



SEQUENCE LISTING

###

(1) GENERAL INFORMATION:

RECEIVED

MAY 12 2000

(i) APPLICANT: EPELBAUM, SABINE URSULA  
FALCO, SAVERIO CARL  
MCDEVITT, RAYMOND ERVIN, III

TECH CENTER 1600/2800

(ii) TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR  
INCREASING THE LYSINE CONTENT OF  
THE SEEDS OF PLANTS

(iii) NUMBER OF SEQUENCES: 132

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
(B) STREET: 1007 MARKET STREET  
(C) CITY: WILMINGTON  
(D) STATE: DELAWARE  
(E) COUNTRY: U.S.A.  
(F) ZIP: 19898

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: DISKETTE, 3.50 INCH  
(B) COMPUTER: IBM PC COMPATIBLE  
(C) OPERATING SYSTEM: MICROSOFT OFFICE 97  
(D) SOFTWARE: MICROSOFT WINDOWS 95

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/824,627  
(B) FILING DATE: MARCH 27, 1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: CHRISTENBURY, LYNNE M.  
(B) REGISTRATION NUMBER: 30,971  
(C) REFERENCE/DOCKET NUMBER: BB-1037-F

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 302-992-5481  
(B) TELEFAX: 302-892-7949  
(C) TELEX: 835420

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1350 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..1350

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GCT GAA ATT GTT GTC TCC AAA TTT GGC GGT ACC AGC GTA GCT GAT	48
Met Ala Glu Ile Val Val Ser Lys Phe Gly Gly Thr Ser Val Ala Asp	
1 5 10 15	
TTT GAC GCC ATG AAC CGC AGC GCT GAT ATT GTG CTT TCT GAT GCC AAC	96
Phe Asp Ala Met Asn Arg Ser Ala Asp Ile Val Leu Ser Asp Ala Asn	
20 25 30	
GTG CGT TTA GTT GTC CTC TCG GCT TCT GCT GGT ATC ACT AAT CTG CTG	144
Val Arg Leu Val Val Leu Ser Ala Ser Ala Gly Ile Thr Asn Leu Leu	
35 40 45	
GTC GCT TTA GCT GAA GGA CTG GAA CCT GGC GAG CGA TTC GAA AAA CTC	192
Val Ala Leu Ala Glu Gly Leu Glu Pro Gly Glu Arg Phe Glu Lys Leu	
50 55 60	
GAC GCT ATC CGC AAC ATC CAG TTT GCC ATT CTG GAA CGT CTG CGT TAC	240
Asp Ala Ile Arg Asn Ile Gln Phe Ala Ile Leu Glu Arg Leu Arg Tyr	
65 70 75 80	
CCG AAC GTT ATC CGT GAA GAG ATT GAA CGT CTG CTG GAG AAC ATT ACT	288
Pro Asn Val Ile Arg Glu Glu Ile Glu Arg Leu Leu Glu Asn Ile Thr	
85 90 95	
GTT CTG GCA GAA GCG GCG GCG CTG GCA ACG TCT CCG GCG CTG ACA GAT	336
Val Leu Ala Glu Ala Ala Ala Leu Ala Thr Ser Pro Ala Leu Thr Asp	
100 105 110	
GAG CTG GTC AGC CAC GGC GAG CTG ATG TCG ACC CTG CTG TTT GTT GAG	384
Glu Leu Val Ser His Gly Glu Leu Met Ser Thr Leu Leu Phe Val Glu	
115 120 125	
ATC CTG CGC GAA CGC GAT GTT CAG GCA CAG TGG TTT GAT GTA CGT AAA	432
Ile Leu Arg Glu Arg Asp Val Gln Ala Gln Trp Phe Asp Val Arg Lys	
130 135 140	
GTG ATG CGT ACC AAC GAC CGA TTT GGT CGT GCA GAG CCA GAT ATA GCC	480
Val Met Arg Thr Asn Asp Arg Phe Gly Arg Ala Glu Pro Asp Ile Ala	
145 150 155 160	

RECEIVED

MAY 08 2000

TECH CENTER 1600/2900

GCG	CTG	GCG	GAA	CTG	GCC	GCG	CTG	CAG	CTG	CTC	CCA	CGT	CTC	AAT	GAA	528
Ala	Leu	Ala	Glu	Leu	Ala	Ala	Leu	Gln	Leu	Leu	Pro	Arg	Leu	Asn	Glu	
			165						170					175		
GGC	TTA	GTG	ATC	ACC	CAG	GGA	TTT	ATC	GGT	AGC	GAA	AAT	AAA	GGT	CGT	576
Gly	Leu	Val	Ile	Thr	Gln	Gly	Phe	Ile	Gly	Ser	Glu	Asn	Lys	Gly	Arg	
			180					185					190			
ACA	ACG	ACG	CTT	GGC	CGT	GGA	GGC	AGC	GAT	TAT	ACG	GCA	GCC	TTG	CTG	624
Thr	Thr	Thr	Leu	Gly	Arg	Gly	Gly	Ser	Asp	Tyr	Thr	Ala	Ala	Leu	Leu	
		195					200					205				
GCG	GAG	GCT	TTA	CAC	GCA	TCT	CGT	GTT	GAT	ATC	TGG	ACC	GAC	GTC	CCG	672
Ala	Glu	Ala	Leu	His	Ala	Ser	Arg	Val	Asp	Ile	Trp	Thr	Asp	Val	Pro	
	210					215					220					
GGC	ATC	TAC	ACC	ACC	GAT	CCA	CGC	GTA	GTT	TCC	GCA	GCA	AAA	CGC	ATT	720
Gly	Ile	Tyr	Thr	Thr	Asp	Pro	Arg	Val	Val	Ser	Ala	Ala	Lys	Arg	Ile	
225					230					235					240	
GAT	GAA	ATC	GCG	TTT	GCC	GAA	GCG	GCA	GAG	ATG	GCA	ACT	TTT	GGT	GCA	768
Asp	Glu	Ile	Ala	Phe	Ala	Glu	Ala	Ala	Glu	Met	Ala	Thr	Phe	Gly	Ala	
			245						250					255		
AAA	GTA	CTG	CAT	CCG	GCA	ACG	TTG	CTA	CCC	GCA	GTA	CGC	AGC	GAT	ATC	816
Lys	Val	Leu	His	Pro	Ala	Thr	Leu	Leu	Pro	Ala	Val	Arg	Ser	Asp	Ile	
			260					265					270			
CCG	GTC	TTT	GTC	GGC	TCC	AGC	AAA	GAC	CCA	CGC	GCA	GGT	GGT	ACG	CTG	864
Pro	Val	Phe	Val	Gly	Ser	Ser	Lys	Asp	Pro	Arg	Ala	Gly	Gly	Thr	Leu	
		275					280					285				
GTG	TGC	AAT	AAA	ACT	GAA	AAT	CCG	CCG	CTG	TTC	CGC	GCT	CTG	GCG	CTT	912
Val	Cys	Asn	Lys	Thr	Glu	Asn	Pro	Pro	Leu	Phe	Arg	Ala	Leu	Ala	Leu	
	290					295					300					
CGT	CGC	AAT	CAG	ACT	CTG	CTC	ACT	TTG	CAC	AGC	CTG	AAT	ATG	CTG	CAT	960
Arg	Arg	Asn	Gln	Thr	Leu	Leu	Thr	Leu	His	Ser	Leu	Asn	Met	Leu	His	
305					310					315				320		
TCT	CGC	GGT	TTC	CTC	GCG	GAA	GTT	TTC	GGC	ATC	CTC	GCG	CGG	CAT	AAT	1008
Ser	Arg	Gly	Phe	Leu	Ala	Glu	Val	Phe	Gly	Ile	Leu	Ala	Arg	His	Asn	
			325						330					335		
ATT	TCG	GTA	GAC	TTA	ATC	ACC	ACG	TCA	GAA	GTG	AGC	GTG	GCA	TTA	ACC	1056
Ile	Ser	Val	Asp	Leu	Ile	Thr	Thr	Ser	Glu	Val	Ser	Val	Ala	Leu	Thr	
			340					345					350			
CTT	GAT	ACC	ACC	GGT	TCA	ACC	TCC	ACT	GGC	GAT	ACG	TTG	CTG	ACG	CAA	1104
Leu	Asp	Thr	Thr	Gly	Ser	Thr	Ser	Thr	Gly	Asp	Thr	Leu	Leu	Thr	Gln	
		355					360					365				
TCT	CTG	CTG	ATG	GAG	CTT	TCC	GCA	CTG	TGT	CGG	GTG	GAG	GTG	GAA	GAA	1152
Ser	Leu	Leu	Met	Glu	Leu	Ser	Ala	Leu	Cys	Arg	Val	Glu	Val	Glu	Glu	
	370					375					380					

GGT	CTG	GCG	CTG	GTC	GCG	TTG	ATT	GGC	AAT	GAC	CTG	TCA	AAA	GCC	TGC	1200
Gly	Leu	Ala	Leu	Val	Ala	Leu	Ile	Gly	Asn	Asp	Leu	Ser	Lys	Ala	Cys	
385					390				395					400		
GCC	GTT	GGC	AAA	GAG	GTA	TTC	GGC	GTA	CTG	GAA	CCG	TTC	AAC	ATT	CGC	1248
Ala	Val	Gly	Lys	Glu	Val	Phe	Gly	Val	Leu	Glu	Pro	Phe	Asn	Ile	Arg	
				405					410					415		
ATG	ATT	TGT	TAT	GGC	GCA	TCC	AGC	CAT	AAC	CTG	TGC	TTC	CTG	GTG	CCC	1296
Met	Ile	Cys	Tyr	Gly	Ala	Ser	Ser	His	Asn	Leu	Cys	Phe	Leu	Val	Pro	
			420					425					430			
GGC	GAA	GAT	GCC	GAG	CAG	GTG	GTG	CAA	AAA	CTG	CAT	AGT	AAT	TTG	TTT	1344
Gly	Glu	Asp	Ala	Glu	Gln	Val	Val	Gln	Lys	Leu	His	Ser	Asn	Leu	Phe	
		435				440						445				
GAG	TAA															1350
Glu	*															
450																

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GATCCATGGC TGAAATTGTT GTCTCCAAAT TTGGCG 36

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTACCGCCAA ATTTGGAGAC AACAAATTCA GCCATG 36

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 48 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCGGGCCAT GGCTACAGGT TTAACAGCTA AGACCGGAGT AGAGCACT

48

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATATCGAAT TCTCATTATA GAACTCCAGC TTTTTTC

37

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 917 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 3..911

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CC ATG GCT ACA GGT TTA ACA GCT AAG ACC GGA GTA GAG CAC TTC GGC	47
Met Ala Thr Gly Leu Thr Ala Lys Thr Gly Val Glu His Phe Gly	
1 5 10 15	
ACC GTT GGA GTA GCA ATG GTT ACT CCA TTC ACG GAA TCC GGA GAC ATC	95
Thr Val Gly Val Ala Met Val Thr Pro Phe Thr Glu Ser Gly Asp Ile	
20 25 30	
GAT ATC GCT GCT GGC CGC GAA GTC GCG GCT TAT TTG GTT GAT AAG GGC	143
Asp Ile Ala Ala Gly Arg Glu Val Ala Ala Tyr Leu Val Asp Lys Gly	
35 40 45	
TTG GAT TCT TTG GTT CTC GCG GGC ACC ACT GGT GAA TCC CCA ACG ACA	191
Leu Asp Ser Leu Val Leu Ala Gly Thr Thr Gly Glu Ser Pro Thr Thr	
50 55 60	
ACC GCC GCT GAA AAA CTA GAA CTG CTC AAG GCC GTT CGT GAG GAA GTT	239
Thr Ala Ala Glu Lys Leu Glu Leu Leu Lys Ala Val Arg Glu Glu Val	
65 70 75	
GGG GAT CGG GCG AAG CTC ATC GCC GGT GTC GGA ACC AAC AAC ACG CGG	287
Gly Asp Arg Ala Lys Leu Ile Ala Gly Val Gly Thr Asn Asn Thr Arg	
80 85 90 95	

ACA	TCT	GTG	GAA	CTT	GCG	GAA	GCT	GCT	GCT	TCT	GCT	GGC	GCA	GAC	GGC	335
Thr	Ser	Val	Glu	Leu	Ala	Glu	Ala	Ala	Ala	Ser	Ala	Gly	Ala	Asp	Gly	
				100					105					110		
CTT	TTA	GTT	GTA	ACT	CCT	TAT	TAC	TCC	AAG	CCG	AGC	CAA	GAG	GGA	TTG	383
Leu	Leu	Val	Val	Thr	Pro	Tyr	Tyr	Ser	Lys	Pro	Ser	Gln	Glu	Gly	Leu	
			115					120					125			
CTG	GCG	CAC	TTC	GGT	GCA	ATT	GCT	GCA	GCA	ACA	GAG	GTT	CCA	ATT	TGT	431
Leu	Ala	His	Phe	Gly	Ala	Ile	Ala	Ala	Ala	Thr	Glu	Val	Pro	Ile	Cys	
		130					135					140				
CTC	TAT	GAC	ATT	CCT	GGT	CGG	TCA	GGT	ATT	CCA	ATT	GAG	TCT	GAT	ACC	479
Leu	Tyr	Asp	Ile	Pro	Gly	Arg	Ser	Gly	Ile	Pro	Ile	Glu	Ser	Asp	Thr	
	145					150					155					
ATG	AGA	CGC	CTG	AGT	GAA	TTA	CCT	ACG	ATT	TTG	GCG	GTC	AAG	GAC	GCC	527
Met	Arg	Arg	Leu	Ser	Glu	Leu	Pro	Thr	Ile	Leu	Ala	Val	Lys	Asp	Ala	
160					165					170				175		
AAG	GGT	GAC	CTC	GTT	GCA	GCC	ACG	TCA	TTG	ATC	AAA	GAA	ACG	GGA	CTT	575
Lys	Gly	Asp	Leu	Val	Ala	Ala	Thr	Ser	Leu	Ile	Lys	Glu	Thr	Gly	Leu	
			180						185					190		
GCC	TGG	TAT	TCA	GGC	GAT	GAC	CCA	CTA	AAC	CTT	GTT	TGG	CTT	GCT	TTG	623
Ala	Trp	Tyr	Ser	Gly	Asp	Asp	Pro	Leu	Asn	Leu	Val	Trp	Leu	Ala	Leu	
			195					200					205			
GGC	GGA	TCA	GGT	TTC	ATT	TCC	GTA	ATT	GGA	CAT	GCA	GCC	CCC	ACA	GCA	671
Gly	Gly	Ser	Gly	Phe	Ile	Ser	Val	Ile	Gly	His	Ala	Ala	Pro	Thr	Ala	
		210					215					220				
TTA	CGT	GAG	TTG	TAC	ACA	AGC	TTC	GAG	GAA	GGC	GAC	CTC	GTC	CGT	GCG	719
Leu	Arg	Glu	Leu	Tyr	Thr	Ser	Phe	Glu	Glu	Gly	Asp	Leu	Val	Arg	Ala	
	225					230					235					
CGG	GAA	ATC	AAC	GCC	AAA	CTA	TCA	CCG	CTG	GTA	GCT	GCC	CAA	GGT	CGC	767
Arg	Glu	Ile	Asn	Ala	Lys	Leu	Ser	Pro	Leu	Val	Ala	Ala	Gln	Gly	Arg	
240					245				250					255		
TTG	GGT	GGA	GTC	AGC	TTG	GCA	AAA	GCT	GCT	CTG	CGT	CTG	CAG	GGC	ATC	815
Leu	Gly	Gly	Val	Ser	Leu	Ala	Lys	Ala	Ala	Leu	Arg	Leu	Gln	Gly	Ile	
				260				265					270			
AAC	GTA	GGA	GAT	CCT	CGA	CTT	CCA	ATT	ATG	GCT	CCA	AAT	GAG	CAG	GAA	863
Asn	Val	Gly	Asp	Pro	Arg	Leu	Pro	Ile	Met	Ala	Pro	Asn	Glu	Gln	Glu	
			275					280					285			
CTT	GAG	GCT	CTC	CGA	GAA	GAC	ATG	AAA	AAA	GCT	GGA	GTT	CTA	TAA	TGAGAATTC	917
Leu	Glu	Ala	Leu	Arg	Glu	Asp	Met	Lys	Lys	Ala	Gly	Val	Leu	*		
		290					295					300				

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTCCCGTGA CCATGGGCCA TC

22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CATGGCTGGC TTCCCCACGA GGAAGACCAA CAATGACATT ACCTCCATTG CTAGCAACGG 60

TGGAAGAGTA CAATG

75

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATGCATTGT ACTCTTCCAC CGTTGCTAGC AATGGAGGTA ATGTCATTGT TGGTCTTCCT 60

CGTGGGGAAG CCAGC

75

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 90 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATGGCTTCC TCAATGATCT CCTCCCCAGC TGTTACCACC GTCAACCGTG CCGGTGCCGG 60

CATGGTTGCT CCATTCACCG GCCTCAAAG

90

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 90 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CATGCTTTTG AGGCCGGTGA ATGGAGCAAC CATGCCGGCA CCGGCACGGT TGACGGTGGT 60  
AACAGCTGGG GAGGAGATCA TTGAGGAAGC 90

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGGTTTGCT GTAATAGGTA CCA 23

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCTTGGTAC CTATTACAGC AAACCGGCAT G 31

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCTTCCTCAA TGATCTCCTC CCCAGCT 27



(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CATTGTACTC TTCCACCGTT GCTAGCAA

28

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..20
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
70"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTGACTCGCT GCGCTCGGTC

20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..24
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
71"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TATTTTCTCC TTACGCATCT GTGC

24

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..27
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
78"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCATCGATA GGCGACCACA CCCGTCC

27

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..27
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
79"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AATATCGATG CCACGATGCG TCCGGCG

27

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..55

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
81"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CATGGAGGAG AAGATGAAGG CGATGGAAGA GAAGATGAAG GCGTGATAGG TACCG 55

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..55  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
80"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AATTCGGTAC CTATCACGCC TTCATCTTCT CTTCCATCGC CTTTCATCTTC TCCTC 55

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: Protein  
(B) LOCATION: 1..14  
(D) OTHER INFORMATION: /label= name  
/note= "base gene  
[(SSP5)2]"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
1 5 10

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..21  
(D) OTHER INFORMATION: /product=  
"synthetic  
oligonucleotide"  
/standard\_name= "SM  
84"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATGGAGGAG AAGATGAAGG C

21

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..21  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
85"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ATCGCCTTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..21  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
82"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GATGGAGGAG AAGCTGAAGG C

21

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..21
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
83"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATCGCCTTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Glu Glu Lys Leu Lys Ala  
1 5

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Glu Glu Lys Met Lys Ala  
1 5

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 160 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
  - (B) STRAIN: E. coli
  - (G) CELL TYPE: DH5 alpha
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: C15
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..151
  - (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.7.7.7.7.5"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG      46
  Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met
    1             5             10             15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG    94
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu
      20             25             30

AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met
      35             40             45

AAG GCG TGATAGGTAC CG                                           160
Lys Ala
    50
```

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 49 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu
  1             5             10             15
```

Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30

Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45

Ala

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 160 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
  - (B) STRAIN: E. coli
  - (G) CELL TYPE: DH5 alpha
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: C20
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..151
  - (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.7.7.7.7.7.5"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30

AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142  
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met  
35 40 45

AAG GCG TGATAGGTAC CG 160  
Lys Ala  
50

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 49 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15

Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30

Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45

Ala

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 139 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: E. coli
- (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:

- (B) CLONE: C30

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..130
- (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "5.7.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30









(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
1 5 10 15

GAG GAG AAG ATG AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95  
Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

CG 97

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15

Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
86"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GATGGAGGAG AAGCTGAAGA A 21

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
87"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATCTTCTTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
88"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GATGGAGGAG AAGCTGAAGT G

21

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
89"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATCCAATTCA GCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
90"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GATGGAGGAG AAGATGAAGA A

21

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
91"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATCTTCTTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..21

(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
92"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GATGGAGGAG AAGATGAAGT G

21

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..21  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
93"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ATCCAATTCA TCTTCTCCTC C

21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Met Glu Glu Lys Leu Lys Lys

1

5

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Glu Glu Lys Leu Lys Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met Glu Glu Lys Met Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Met Glu Glu Lys Met Lys Trp  
1 5

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 160 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:  
(B) STRAIN: E. coli  
(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
(B) CLONE: 82-4

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 2..151  
(D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"



/gene= "ssp"  
/standard\_name=  
"7.7.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

C ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG 46  
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met  
1 5 10 15  
GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG 94  
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu  
20 25 30  
AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAA GAG AAG ATG 142  
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met  
35 40 45  
AAG GCG TGATAGGTAC CG 160  
Lys Ala  
50

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 49 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu  
1 5 10 15  
Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys  
20 25 30  
Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45  
Ala

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 97 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:  
(B) STRAIN: E. coli  
(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
(B) CLONE: 84-H3

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 2..88  
(D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.5.5.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46  
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met  
1 5 10 15

GAG GAG AAG ATG AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95  
Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

CG 97

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15

Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20 25

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 97 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:  
(B) STRAIN: E. coli  
(G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
(B) CLONE: 86-H23

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 2..88  
 (D) OTHER INFORMATION: /function= "synthetic  
 storage protein  
 /product= "protein"  
 /gene= "ssp"  
 /standard\_name=  
 "5.8.8.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG CTG AAG AAG ATG 46  
 Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Lys Met  
 1 5 10 15  
 GAG GAG AAG CTG AAG AAG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95  
 Glu Glu Lys Leu Lys Lys Met Glu Glu Lys Met Lys Ala  
 20 25  
 CG 97

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 28 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Lys Met Glu  
 1 5 10 15  
 Glu Lys Leu Lys Lys Met Glu Glu Lys Met Lys Ala  
 20 25

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 112 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:  
 (B) STRAIN: E. coli  
 (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: 88-2

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 2..103  
 (D) OTHER INFORMATION: /function= "synthetic  
 storage protein  
 /product= "protein"  
 /gene= "ssp"  
 /standard\_name=  
 "5.9.9.9.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

C ATG GAG GAG AAG ATG AAG GCG AAG AAG CTG AAG TGG ATG GAG GAG 46  
 Met Glu Glu Lys Met Lys Ala Lys Lys Leu Lys Trp Met Glu Glu  
 1 5 10 15  
 AAG CTG AAG TGG ATG GAG GAG AAG CTG AAG TGG ATG GAA GAG AAG ATG 94  
 Lys Leu Lys Trp Met Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Met  
 20 25 30  
 AAG GCG TGATAGGTAC CG 112  
 Lys Ala

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 33 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Glu Glu Lys Met Lys Ala Lys Lys Leu Lys Trp Met Glu Glu Lys  
 1 5 10 15  
 Leu Lys Trp Met Glu Glu Lys Leu Lys Trp Met Glu Glu Lys Met Lys  
 20 25 30  
 Ala

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 118 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:  
 (B) STRAIN: E. coli  
 (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: 90-H8

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 2..109  
 (D) OTHER INFORMATION: /function= "synthetic  
 storage protein  
 /product= "protein"  
 /gene= "ssp"  
 /standard\_name=  
 "5.10.10.10.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG AAG ATG 46  
 Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Lys Met  
 1 5 10 15  
 GAG GAG AAG ATG AAG AAG ATG GAG GAG AAG ATG AAG AAG ATG GAA GAG 94  
 Glu Glu Lys Met Lys Lys Met Glu Glu Lys Met Lys Lys Met Glu Glu  
 20 25 30  
 AAG ATG AAG GCG TGATAGGTAC CG 118  
 Lys Met Lys Ala  
 35

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Lys Met Glu  
 1 5 10 15  
 Glu Lys Met Lys Lys Met Glu Glu Lys Met Lys Lys Met Glu Glu Lys  
 20 25 30  
 Met Lys Ala  
 35

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 97 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:  
 (B) STRAIN: E. coli  
 (G) CELL TYPE: DH5 alpha

(vii) IMMEDIATE SOURCE:  
(B) CLONE: 92-2

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 2..88  
(D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.11.11.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```
C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG TGG ATG 46
  Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Trp Met
    1             5             10             15

GAG GAG AAG ATG AAG TGG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC 95
Glu Glu Lys Met Lys Trp Met Glu Glu Lys Met Lys Ala
    20             25

CG 97
```

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Trp Met Glu
  1             5             10             15

Glu Lys Met Lys Trp Met Glu Glu Lys Met Lys Ala
    20             25
```

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 84 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1..84  
(D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"

/standard\_name= "SM  
96"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GATGGAGGAA AAGATGAAGG CGATGGAGGA GAAAATGAAA GCTATGGAGG AAAAGATGAA 60  
AGCGATGGAG GAGAAAATGA AGGC 84

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..84
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
97"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATCGCCTTCA TTTTCTCCTC CATCGCTTTC ATCTTTTCCT CCATAGCTTT CATTTTCTCC 60  
TCCATCGCCT TCATCTTTTC CTCC 84

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..28
- (D) OTHER INFORMATION: /label= name  
/note= "(SSP 5)4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu  
1                   5                   10                   15  
Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala  
20                   25

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..84
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
98"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GATGGAGGAA AAGCTGAAAG CGATGGAGGA GAAACTCAAG GCTATGGAAG AAAAGCTTAA 60  
AGCGATGGAG GAGAAACTGA AGGC 84

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 84 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..84
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
99"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

ATCGCCTTCA GTTTCTCCTC CTACGCTTTA AGCTTTTCTT CCATAGCCTT GAGTTTCTCC 60  
TCCATCGCTT TCAGCTTTTC CTCC 84

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein









(vii) IMMEDIATE SOURCE:

(B) CLONE: 5-1

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..172

(D) OTHER INFORMATION: /function= "synthetic  
storage protein  
/product= "protein"  
/gene= "ssp"  
/standard\_name=  
"5.5.5.7.7.7.5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```
C ATG GAG GAG AAG ATG AAG GCG ATG GAG GAG AAG ATG AAG GCG ATG 46
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met
1 5 10 15

GAG GAG AAG ATG AAG GCG ATG GAG GAA AAG CTG AAA GCG ATG GAG GAG 94
Glu Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu
20 25 30

AAA CTC AAG GCT ATG GAA GAA AAG CTT AAA GCG ATG GAG GAG AAA CTG 142
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu
35 40 45

AAG GCC ATG GAA GAG AAG ATG AAG GCG TGATAG 175
Lys Ala Met Glu Glu Lys Met Lys Ala
50 55
```

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```
Met Glu Glu Lys Met Lys Ala Met Glu Glu Lys Met Lys Ala Met Glu
1 5 10 15

Glu Lys Met Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys
20 25 30

Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys
35 40 45

Ala Met Glu Glu Lys Met Lys Ala
50 55
```

(2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 187 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
  - (B) STRAIN: E. coli
  - (G) CELL TYPE: DH5 alpha
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 3..173
  - (D) OTHER INFORMATION: /function= "synthetic storage protein"  
/product= "protein"  
/gene= "ssp"  
/standard\_name= "SSP-3-5"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```
CC ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG      47
  Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met
    1                5                10                15

GAG GAG AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAG      95
Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu
                20                25                30

AAG CTG AAG GCG ATG GAG GAG AAG CTG AAG GCG ATG GAG GAA AAG ATG     143
Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met
                35                40                45

AAG GCG ATG GAA GAG AAG ATG AAG GCG TGATAGGTAC CGAATTC             187
Lys Ala Met Glu Glu Lys Met Lys Ala
    50                55
```

(2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 56 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```
Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu
  1                5                10                15

Glu Lys Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys
    20                25                30
```

Leu Lys Ala Met Glu Glu Lys Leu Lys Ala Met Glu Glu Lys Met Lys  
35 40 45

Ala Met Glu Glu Lys Met Lys Ala  
50 55

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..61
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
107"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

CATGGAGGAG AAGATGAAAA AGCTCGAAGA GAAGATGAAG GTCATGAAGT GATAGGTACC 60  
G 61

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..61
- (D) OTHER INFORMATION: /product= "synthetic  
ligonucleotide"  
/standard\_name= "SM  
106"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

AATTCGGTAC CTATCACTTC ATGACCTTCA TCTTCTCTTC GAGCTTTTTC ATCTTCTCCT 60  
C 61



- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..63
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
111"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

AGCTTCTTCA TCTTTTCCTC AAGCCATTTT ATCTTGTCCT CCATAGCCTT CATCTTTTCT 60  
TCC 63

(2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Lys  
35

(2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys Ala Met Glu  
1 5 10 15  
Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Lys  
35



(2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..62
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
112"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GCTCGAAGAA AGATGAAGGC AATGGAAGAC AAAATGAAGT GGCTTGAGGA GAAAATGAAG 60  
AA 62

(2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..62
  - (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
113"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

AGCTTCTTCA TTTTCTCCTC AAGCCACTTC ATTTTGTCTT CCATTGCCTT CATCTTTCTT 60  
CG 62

(2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Met Glu Glu Lys Met Lys Lys Leu Lys Glu Glu Met Ala Lys Met Lys  
1 5 10 15  
Asp Glu Met Trp Lys Leu Lys Glu Glu Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Lys  
35

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 63 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..63
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
114"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

GCTCAAGGAG GAAATGGCTA AGATGAAAGA CGAAATCTGG AACTGAAAG AGGAAATGAA 60  
GAA 63

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 63 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..63
- (D) OTHER INFORMATION: /product= "synthetic  
oligonucleotide"  
/standard\_name= "SM  
115"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

AGCTTCTTCA TTTCTCTTT CAGTTTCCAC ATTTGCTTT TCATCTTAGC CATTTCCTCC 60  
TTG 63

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Met Glu Glu Lys Met Lys Lys Leu Lys Glu Glu Met Ala Lys Met Lys  
1 5 10 15  
Asp Glu Met Trp Lys Leu Lys Glu Glu Met Lys Lys Leu Glu Glu Lys  
20 25 30  
Met Lys Val Met Glu Glu Lys Met Lys Lys Leu Glu Glu Lys Met Lys  
35 40 45  
Ala Met Glu Asp Lys Met Lys Trp Leu Glu Glu Lys Met Lys Lys Leu  
50 55 60  
Glu Glu Lys Met Lys Val Met Glu Glu Lys Met Lys Lys Leu Glu Glu  
65 70 75 80  
Lys Met Lys Ala Met Glu Asp Lys Met Lys Trp Leu Glu Glu Lys Met  
85 90 95  
Lys Lys Leu Glu Glu Lys Met Lys Val Met Lys  
100 105

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 839 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGATCCCCCG GGCTGCAGGA ATTCTACGTA CCATATAGTA AGACTTTGTA TATAAGACGT 60  
CACCTCTTAC GTGCATGGTT ATATGTGACA TGTGCAGTGA CGTTGTACCA TATAGTAAGA 120  
CTTTGTATAT AAGACGTCAC CTCTTACGTG CATGGTTATA TGTGACATGT GCAGTGACGT 180  
TAACCGCACC CTCCTTCCCG TCGTTTCCCA TCTCTTCCTC CTTTAGAGCT ACCACTATAT 240  
AAATCAGGGC TCATTTTCTC GTCCTCACA GGCTCATCAG CACCCCGGCA GTGCCACCCC 300  
GACTCCCTGC ACCTGCCATG GGTACGCTAG CCCGGGAGAT CTGACAAAGC AGCATTAGTC 360  
CGTTGATCGG TGGAAGACCA CTCGTCAGTG TTGAGTTGAA TGTTTGATCA ATAAAATACG 420

GCAATGCTGT AAGGGTTGTT TTTTATGCCA TTGATAATAC ACTGTACTGT TCAGTTGTTG 480  
 AACTCTATTT CTTAGCCATG CCAGTGCTTT TCTTATTTTG AATAACATTA CAGCAAAAAG 540  
 TTGAAAGACA AAAAAANNNN NCCCCGAACA GAGTGCTTTG GGTCCCAAGC TTCTTTAGAC 600  
 TGTGTTCCGC GTTCCCCCTA AATTTCTCCC CTATATCTCA CTCACTTGTC ACATCAGCGT 660  
 TCTCTTTCCC CTATATCTCC ACGCTCTACA GCAGTCCAC CTATATCAAA CCTCTATACC 720  
 CCACCACAAC AATATTATAT ACTTTCATCT TCACCTAACT CATGTACCTT CCAATTTTTT 780  
 TCTACTAATA ATTATTTACG TGCACAGAAA CTTAGGCAAG GGAGAGAGAG AGCGGTACC 839

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CTAGAAGCCT CGGCAACGTC AGCAACGGCG GAAGAATCCG GTG 43

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

CATGCACCGG ATTCTCCGC CGTTGCTGAC GTTGCCGAGG CTT 43

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GATCCCATGG CGCCCCTTAA GTCCACCGCC AGCCTCCCCG TCGCCGCCG CTCCT 55

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CTAGAGGAGC GCGGGCGAC GGGGAGGCTG GCGGTGGACT TAAGGGGCGC CATGG 55

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CATGGCGCCC ACCGTGATGA TGGCCTCGTC GGCCACCGCC GTCGCTCCGT TCCAGGGGC 59

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

TTAAGCCCCT GGAACGGAGC GACGGCGGTG GCCGACGAGG CCATCATCAC GGTGGGCGC 59

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GCGCCCACCG TGATGA

16

(2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CACCGGATTC TTCCGC

16

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 372 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

GTAAGATTGG TAAAGTCCAG CAAGAAAATG AGATAAAAGA GAAGCCTGAA ATGACGAAAA 60  
AATCAGGTGT TTTGATTCTT GGTGCTGGAC GTGTGTNTCG CCCAGCTGCT GATTTCCTAG 120  
CTTCAGTTAG AACCATTTCG TCACAGCAAT GGTACAAAAC ATATTTCCGGA GCAGACTCTG 180  
AAGAGAAAAC AGATGTTTCAT GTGATTGTCG CGTCTCTGTA TCTTAAGGAT GCCAAAGAGA 240  
CGGTTGAAGG TATTTTCAGAT GTAGAAGCAG TTCGGCTAGA TGTATCTGAT AGTGAAAGTC 300  
TCCTTAAGTA TGTTTCTCAG GTTGATGTTG TCCTAAGTTT ATTACCTGCA AGTTGTCATG 360  
CTTGTTGTAG CA 372

(2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 323 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GGAAGCACAC TGCGACTCTT TTGGAATTCG GGGACATCAA GAATGGACAA ACAACAACCG 60  
CTATGGCCAA GACTGTTGGG ATCCCTGCAG CCATTGGAGC TCTGCTGTTA ATTGAAGACA 120  
AGATCAAGAC AAGAGGAGTC TTAAGGCCTC TCGAAGCAGA GGTGTATTTG CCAGCTTTGG 180

ATATATTGCA AGCATATGGT ATAAAGCTGA TGGAGAAGGC AGAATGATCA AAGAACTCTG 240  
TATATTGTTT CTNCTATAA CTTGGAGTTG GAGACAAAGC TGAAGGAGNC AGNGCCATTA 300  
GACCAGCAAA AAAAGGAGGA GGA 323

(2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 123 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Lys Ile Gly Lys Val Gln Gln Glu Asn Glu Ile Lys Glu Lys Pro Glu  
1 5 10 15  
Met Thr Lys Lys Ser Gly Val Leu Ile Leu Gly Ala Gly Arg Val Xaa  
20 25 30  
Arg Pro Ala Ala Asp Phe Leu Ala Ser Val Arg Thr Ile Ser Ser Gln  
35 40 45  
Gln Trp Tyr Lys Thr Tyr Phe Gly Ala Asp Ser Glu Glu Lys Thr Asp  
50 55 60  
Val His Val Ile Val Ala Ser Leu Tyr Leu Lys Asp Ala Lys Glu Thr  
65 70 75 80  
Val Glu Gly Ile Ser Asp Val Glu Ala Val Arg Leu Asp Val Ser Asp  
85 90 95  
Ser Glu Ser Leu Leu Lys Tyr Val Ser Gln Val Asp Val Val Leu Ser  
100 105 110  
Leu Leu Pro Ala Ser Cys His Ala Cys Cys Ser  
115 120

(2) INFORMATION FOR SEQ ID NO:105:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 74 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

Lys His Thr Ala Thr Leu Leu Glu Phe Gly Asp Ile Lys Asn Gly Gln  
1 5 10 15

Thr Thr Thr Ala Met Ala Lys Thr Val Gly Ile Pro Ala Ala Ile Gly  
20 25 30

Ala Leu Leu Leu Ile Glu Asp Lys Ile Lys Thr Arg Gly Val Leu Arg  
35 40 45

Pro Leu Glu Ala Glu Val Tyr Leu Pro Ala Leu Asp Ile Leu Gln Ala  
50 55 60

Tyr Gly Ile Lys Leu Met Glu Lys Ala Glu  
65 70

(2) INFORMATION FOR SEQ ID NO:106:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

ATTCCCCATG GTTTCGCCGA CGAAT 25

(2) INFORMATION FOR SEQ ID NO:107:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

CTCTCGGTAC CTAGTACCTA CTGATCAAC 29

(2) INFORMATION FOR SEQ ID NO:108:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

AGAGAAGCCT GAAATGACGA AAAA 24



(2) INFORMATION FOR SEQ ID NO:109:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

GTCTTGGCCA TAGCGTGTG TGTT

24

(2) INFORMATION FOR SEQ ID NO:110:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8160 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

TCTAGATGCA CATTCAACTC GAGGTTGTTG CATGATGTTT CATTACCAA AAAAATCATA 60  
GTCAAATTAT GTAAGCAAAT GATATTACAG AAAAGTTTTA CTAGAGAGTT TCAGATTTAC 120  
ACATGCACAA CGTTAAAAA AATAGCAGAA AAAAGAAAGA AGAAAAGTTC TTTATTTGTG 180  
AGAAAAATGT ATGAAAAAA AAGAGATGGG TGTA AAAAGC AAAAGGATAG GACCACTGTT 240  
ACTTTGTAGC CTCGTTGAGG AATCTCTTCT CGCATCTCGA CTTTTGTGCC ATTGCAAAGT 300  
CAATGCCCGAG AACTTGTTCC CAGGCCATCT CCAATTAACT ACGTCTATTT AATTAAACTT 360  
TTAAAAGAAA ACCTAATAAA TTAACA AAAAGCCGT CAACGAAATC TAAGCTTGCA 420  
GCGATATCGA TGAAGTATA CCAAACAAT GTTCAAGTTT CACTTTCAA TTTGTTTTTC 480  
TTGAAATAGT TTATTGGGTA AGGCCCATAG ATATTTCATA AGAAGAACAC TTGTCGAGGT 540  
TGAATCGTAT GTCTGCCAC CGCGGCCAT GCATCCTCTG TTGGTAGCAT AATCGTTTTA 600  
GGCCATACTA TTGTTTCGTAC AACTGATTT TGAAGTCACC TTTGTGCACT CCTTAATTCC 660  
TAAATTGAAG AAGCTTGTTT TCATTCTTCT TTGGGTTACA AATGCCAAGG CAAAAGGAAC 720  
TTGGGCCAAA TTAAGACAAC AACTCAAGCC CACTCTCTGC AAATAATACT TGGGAATTTT 780  
TACTAAAACG GTGCGTTTCA TCCAAGAATC TATTAATATC CCTAACTTGA AATCATCATA 840  
TACGTAACCC AACATATTAA AGAGTTAATA ATGTTAAAAA AAGTCTCAGA AGAGAGAGAC 900  
GTAGAGAACA CGAAAGTGG TAACTGGTAA GCGTCGTCAT CGAGGATATA GTAGCTACGT 960

GAGCAAACGT	CTTCACTCAT	CTCTGTCTAT	TTCTCTTCGA	ATACACGTAA	TACATTTTCG	1020
ATTGGATTGA	TCCTCCCTCG	GTCCTATCCA	AGTATCCATC	CACGTAAACA	AGAGCTTGTT	1080
CCTTTCTTGT	TTTTTCTTTC	TTTAAATAGT	AAAAATACTT	ATTCATTTG	TTTCGTTTGA	1140
TTTCATTATT	ATTGTCTATG	GCATTATATA	CTATATATAT	TATTTCTACA	ACATTGGCTG	1200
GCTCACGTTG	TTCTCGTGTA	TACAACAAAC	TTAATTAATG	TCTCTCTATT	GCATTAGATA	1260
GTTTCGGAGC	ATATCCATTA	TGTGAAAGCC	ACATTAAGTT	ATAACTAAAA	GTAGTTTTTCG	1320
AAAGAGCTTA	ATTAAGTTAT	GTTCTGTTTC	AAATAAAAAAT	GAACACGAGG	GATTTTTTTTT	1380
TTTTTTGACA	GATCATTATT	AACAAAAATG	ATTACCTGAA	GAAAGGGGAA	AATAATTATA	1440
GCTGATTACA	GATCATTATT	AACAAAAAGA	ATTCTTGTC	CATCATTTCAT	TATAACAAGA	1500
AATATTATAT	TATATTAATT	TAATCTTTTCG	CTAACACGCC	CACAATATAT	TAATCATATA	1560
CGTAATTTAG	CTTATAAAAA	GGACGGAAAG	AGATTATTAC	TGCGCCTAAA	AAACTCACTA	1620
ATTCCAAAGA	AAAAAAAAAAG	CTTGTATTTT	TTCTTGACAA	ACCAGCTCAC	AGGCATTGCA	1680
TGATCAAAC	CATCAGGTAC	GTTTTGATTC	CTTCTTCCAT	AATTTTCCCA	TCTTGAGGAA	1740
TGCAAATTTG	GAGAGCGCTT	TAGCTAAATC	ACTGCCTTCA	TTTTTTCACT	TTGGATTTAA	1800
TAATTTGCAT	TCCTCTCTTC	CTCTCTGCTC	TGTTCTGTTC	TGTTCTGTTC	TGATTTGAGT	1860
TTTCAATTAA	TCGCTCGAGC	AAAAGCTATT	TCTCAACTCG	TTAAATTTCT	GTTCCAGTT	1920
TGTTTCGATTT	TCAACAGTTT	CACATTAAG	TTTGGGTTTT	TGATGTTTGG	TTGATGAAAC	1980
TCGAAATATG	AAATGTTTGT	GAATCTATTC	CAGGGTGTTT	AAAATAAGGG	TTTGTGTTC	2040
ATCTGCAGAG	ATTATATGTT	TTTACATGAA	AGATGAATTC	AAATGGCCAT	GAGGAGGAGA	2100
AGAAGTTGGG	GAATGGAGTT	GTGGGGATTC	TAGCTGAAAC	AGTTAACAAA	TGGGAGAGAC	2160
GAACACCATT	GACGCCATCG	CATTGCGCTC	GCCTTTTACA	CGGTGGGAAA	GACAGAACCG	2220
GCATTTCCCG	CATTGTGGTT	CAGCCATCTG	CTAAGCGTAT	CCATCATGAT	GCCTTGTATG	2280
AAGATGTTGG	GTGTGAAATT	TCTGATGATT	TGTCTGATTG	TGGGCTTATA	CTTGGAAATCA	2340
AACAACCTGA	GGTGTGGGAA	TTTGCATTAA	AAAGAGTTCC	TTTTTTTCTT	CTATATATAT	2400
ATCAGTTTAT	GAGATTTGAT	TCTGTTTGCA	GCTAGAAATG	ATTCTTCCAG	AGAGAGCATA	2460
CGCTTTCTTT	TCACATACTC	ATAAGGCACA	GAAAGAGAAC	ATGCCTTTGT	TGGATAAAGT	2520
ATTACACTTT	TCATTTATCC	TTTTAGTCCT	ATCTAAGATA	CTGAGGAATG	TTGACAAAAG	2580
GGGTATCCAA	TTGCAGATTC	TTTCTGAGAG	AGTGACTTTG	TGTGATTATG	AGCTCATTGT	2640
TGGGGATCAT	GGGAAACGAT	TATTGGCGTT	TGGTAAATAT	GCAGGCAGAG	CTGGTCTTGT	2700

TGACTTCTTA	CACGGACTTG	GACAGCGTAA	GCTCATGTTA	TAATTCTGAT	GATCAGGACA	2760
TGTTTCTGTG	CAGAACAAGA	TGAGATGTAA	TTTTCCATGT	TTGATGCAGG	ATATCTAAGT	2820
CTAGGATACT	CAACACCTTT	CCTCTCGCTC	GGTGCATCGT	ATATGTATTC	CTCATTGGCT	2880
GCTGCAAAAAG	CCGCTGTAAT	TTCTGTTGGT	GAAGAAATTG	CAAGCCAGGG	ACTGCCATTA	2940
GGAACTG GCC	CTCTTG TATT	TGTCTTCACC	GGAACAGGAA	ATGGTATCTT	CTTTAGTTCT	3000
ACTGCGAGTT	CTTTGAATCC	TTCTGCATAT	GTTTCATCTC	ATTAAAAAAT	TTCTCATCCG	3060
CAGTTTCTCT	GGGGGCGCAA	GAAATTTTCA	AGCTTCTTCC	TCACACTTTT	GTTGAACCAA	3120
GCAA ACTTCC	TGAACTATTT	GTAAAAGTAA	GTCACGCTTT	GCTTTTTATT	TGGTTTCAGA	3180
GTTTTGAAGA	TTCTGAAATG	TATATTTCTC	ACAGGACAAA	GGAATTAGTC	AAAATGGGAT	3240
TTCAACAAAG	CGAGTCTATC	AAGTATATGG	TTGTATTATT	ACCAGCCAAG	ACATGGTTGA	3300
ACACAAAGAT	CCATCAAAGT	CATTTCGACAA	AGTAACACTT	ACCTTCTTAG	CTCCTTGGCT	3360
GTGACTTTTG	TTCCACTACG	CTAAAGTAGA	ATACCTATTA	ATTCTTCAAG	CTTATGATGT	3420
TTAGGCCGAC	TATTATGCAC	ACCCGGAACA	TTACAATCCA	GTTTTCCACG	AAAAGATATC	3480
GCCATATACG	TCTGTTCTTG	GTAGATCCTG	ATCACTGTTT	TACCTTTAAA	GCTCAAGAGT	3540
TTACATATAA	GCAAATCCTC	TGTCCACTCC	GTGACTGTGA	CCATCTCATT	TTGGTTAGTT	3600
CCAGTGTGTA	ACCCCTATGA	CTTTCTGTGC	AGTAAACTGT	ATGTACTGGG	AGAAGAGGTT	3660
TCCCTGTCTT	CTGAGCACAA	AACAGCTTCA	AGATTTAACA	AAAAAAGGAC	TCCC ACTAGT	3720
AGGCATATGT	GATATAACTT	GTGACATCGG	TGGCTCCATT	GAATTTGTTA	ACCGAGCTAC	3780
TTTAATCGAT	TCCCCTTTCT	TCAGGTAATA	TATACTTAGG	AAGAGCTTTC	TTTTGAGTCA	3840
TCTACGTTTA	CTATGATGAA	ACTCGTCGAG	CTAAACACTA	TCTCTAGGTT	TAATCCCTCG	3900
AACAATTCAT	ACTACGATGA	CATGGATGGG	GATGGCGTAC	TATGCATGGC	TGTTGACATT	3960
TTACCCACAG	AATTTGCAAA	AGAGGTATGT	ATGAAGGTTA	CAGTTATAGT	ACTTAAGATT	4020
AAATCTAAAG	TTAAAAACCT	TGTATTGAGT	GGGAGTTCTT	GTGTCCTGAA	AAAGGCATCC	4080
CAGCATTTTG	GAGATATTCT	TTCCGGATTT	GTCGGTAGTT	TGGCTTCAAT	GACTGAAATT	4140
TCAGATCTAC	CAGCACATCT	GAAGAGGGCT	TGCATAAGCT	ATAGGGGAGA	ATTGACATCT	4200
TTGTATGAGT	ATATTCCACG	TATGAGGAAG	TCAAATCCAG	AGTATGTTCT	GCTTCGAGCG	4260
TTACTTCATC	TGAAATATTT	AGGCCTCTTC	TCTAAACTAT	GTTTTCATCT	TTACCCACTT	4320
TAACTGCAGA	GAGGCACAAG	ATAATATTAT	CGCCAACGGG	GTTTCCAGCC	AGAGAACATT	4380
CAACATATTG	GTTAGTTTTG	ATGAAGAAAG	TATATATAAC	TAGTTTCCGA	ATCATATGAT	4440

TTAAGCTAAT GAATTAAGAA AATATATAGT TCAAGACTTA TGATTCATAT CTCTATCAAC 4500  
 TTTTTGACCA AAGATTGATA CTTTTTCGAC ATCTGTCACA GCATTTTGTG ATGATTTTGA 4560  
 TTGAGACAAA TCATTTGTAG GTATCTCTGA GCGGACACCT ATTTGATAAG TTTCTGATAA 4620  
 ACGAAGCTCT TGATATGATC GAAGCGGCTG GTGGCTCATT TCATTTGGCT AAATGTGAAC 4680  
 TGGGGCAGAG CGCTGATGCT GAATCGTACT CAGAACTTGA AGTAAGTTTC TTTCTGGATA 4740  
 AAACCTAATC ATTCACATGG AACAACTGTC AAGAGTTTTT AATGTCACGT TTAGGTTCAA 4800  
 TGTCCTTTTC ACTAAGTCTC GTAAGTTTTT AAAACAAGTA AACAACTAC AAGCCAAAAA 4860  
 CATTCTGGCC CCACATTAAC CTATTCCCAC TTGTTAAAGA ACCCATCTTG CATTATCTTG 4920  
 GTAGGTTGGT GCGGATGATA AGAGAGTATT GGATCAAATC ATTGATTCAT TAACTCGGTT 4980  
 AGCTAATCCA AATGAAGATT ATATATCCCC ACATAGAGAA GCAAATAAGA TCTCACTGAA 5040  
 GATTGGTAAA GTCCAGCAAG AAAATGAGAT AAAAGAGAAG CCTGAAATGA CGAAAAAATC 5100  
 AGGCGTTTTG ATTCTTGGTG CTGGACGTGT GTGTCGCCCA GCTGCTGATT TCCTAGCTTC 5160  
 AGTTAGAACC ATTTTCGTCAC AGCAATGGTA CAAAACATAT TTCGGAGCAG ACTCTGAAGA 5220  
 GAAAACAGAT GTTCATGTGA TTGTCGCGTC TCTGTATCTT AAGGATGCCA AAGAGGTAGG 5280  
 AGAAGCCTTT GGGCTTCATC TGAGTAATTC AGTGTATACG ATGAACTATC AATCTTTTAA 5340  
 AGTTTTACTG ATGATCAAAT TTTCCGCAGA CGGTTGAAGG TATTTTCAGAT GTAGAAGCAG 5400  
 TTCGGCTAGA TGTATCTGAT AGTGAAAGTC TCCTTAAGTA TGTTTCTCAG GTATTTTCCT 5460  
 AACTTCTCTG TTCTTAGATC ACCTTTACTT CAAACTCCAC TGTTCAAATC CATGATCTTA 5520  
 TATTTTTTTT TCATTGCACG CAGGTTGATG TTGTCTAAG TTTATTACCT GCAAGTTGTC 5580  
 ATGCTGTTGT AGCAAAGACA TGCATTGAGG TAAATTCCTA ACGTTTAATG CGTTTTCCGA 5640  
 GTGAAGTTAT GAAATTTGCA AATGTTATTC GACATAGAGG TTAAACTTCC TCTGCATAAC 5700  
 ACATTCTTTC AGTAGTTTCC GGTTCCATAA TGTCTCTGTT TCTTCTTTCT GATTCACTCA 5760  
 GCTGAAGAAG CATCTCGTCA CTGCTAGCTA TGTTGATGAT GAAACGTCCA TGTTACATGA 5820  
 GAAGGCTAAG AGTGCTGGGA TAACGATTCT AGGCGAAATG GGAAGTGGACC CTGGAATCGG 5880  
 TATGATATCT CACAACATAG TATCTCTTAA GATCATTGTG TCACTTGATT TAACTTAAGT 5940  
 GCATTTATCT TCAAATATT TCCCGGATAA CTGAGAAGGT GATCCTACAA TGAATCTTTC 6000  
 AGATCACATG ATGGCGATGA AAATGATCAA CGATGCTCAT ATCAAAAAAG GGAAAGTGAA 6060  
 GTCTTTTACC TCTTATTGTG GAGGGCTTCC CTCTCCTGCT GCAGCAAATA ATCCATTAGC 6120  
 ATATAAATTT AGGTACGGTA GTCCTTTACG CCATTAACAT ATTTTGTTTT GTTTAACTCA 6180

TTTAGACATC CTTTCAGAAT TTCGCTTACT CAATTACATC TCGGTATTTT CAGCTGGAAC 6240  
 CCTGCTGGAG CAATTCGAGC TGGTCAAAAC CCCGCCAAAT ACAAAGCAA CGGCGACATA 6300  
 ATACATGTTG ATGGTATGAA AAACAAAATA TGTCTACATG CAGGAGAGGT TGGAGTAGTT 6360  
 TAGCTTCACT ACACATCATT TTTGTTTAAAC CGAGCAATGT AAATCGCAGG GAAGAATCTC 6420  
 TATGATTCCG CGGCAAGATT CCGAGTACCT AATCTTCCAG CTTTTGCATT GGAGTGTCTT 6480  
 CCAAATCGTG ACTCCTTGGT TTACGGGGAA CATTATGGCA TCGAGAGCGA AGCAACAACG 6540  
 ATATTTTCGTG GAACACTCAG ATATGAAGGC ATGAATTCCA TAATCACAAC TCACGACTCA 6600  
 CTTCTCCATA TCTGAAGGCT TAACACTTGT TTTCTTTTGG CTTGTACAGG GTTTAGTATG 6660  
 ATAATGGCAA CACTTTTCGAA ACTTGGATT CTTGACAGTG AAGCAAATCA AGTACTCTCC 6720  
 ACTGGAAAGA GGATTACGTT TGGTGCTCTT TTAAGTAACA TTCTAAATAA GGATGCCGAC 6780  
 AATGAATCAG AGCCCCTAGC GGGAGAAGAA GAGATAAGCA AGAGAATTAT CAAGCTTGGG 6840  
 CATTCCAAGG AGACTGCAGC CAAAGCTGCC AAAACAATTG TGTAAGCTTC TCCATGAAGA 6900  
 TATATAATCT GAATGTTGCA GTGTGATTCC AATTCTTCTA CGAAACTCCT AACCCCAATT 6960  
 CTTTTGTGGT GTCTTAGATT CTTGGGGTTC AACGAAGAGA GGGAGGTTCC ATCACTGTGT 7020  
 AAAAGCGTAT TTGATGCAAC TTGTTACCTA ATGGAAGAGA AACTAGCTTA TTCCGAAAT 7080  
 GAACAGGTCT CTGTTTCATG TGAAAGCATT AGTTTTCTTC TCTCACTTGT ATTTGGTGTT 7140  
 ACTTACTGAC ATAAACTTTG GACAATCTTT TGCATTATGT TTTCAGGACA TGGTGCTTTT 7200  
 GCATCACGAA GTAGAAGTGG AATTCCTTGA AAGCAAACGT ATAGAGAAGC AACTGCGAC 7260  
 TCTTTTGGAA TTCGGGGACA TCAAGAATGG GCAAACAACA ACCGCTATGG CCAAGACTGT 7320  
 TGGGATCCCT GCAGCCATTG GAGCTCTGGT CCTTACTAAG ACTTTGATCA CCACTTTTTT 7380  
 CTGTCTATAT TTCTCTAAAA TGAAAGTTTT AAGCGTTTGT TTTATGATGT TGTGTGTTGC 7440  
 AGCTGTTAAT TGAAGACAAG ATCAAGACAA GAGGAGTCTT AAGGCCTTTC GAAGCAGAGG 7500  
 TGTATTTGCC AGGTAAATTA GAATTCGGCT TCAAAGGAT GTGTGTTGCA GATAAAGACA 7560  
 ATGATGTTGA TTTGTTGTGT GTTTGGGATA TGTGGTGTTA TACATACAGC TTTGGATATA 7620  
 TTGCAAGCAT ATGGTATAAA GCTGATGGAG AAGGCAGAAT GATCAAAGAA CTCTGTATAT 7680  
 TGTTTCTCTC TATAACTTGG AGTTGGAGAC AAAGCTGAAG AAGACAGAGA CATTAGACCA 7740  
 GCAAAAAAAG AAGAAGAAGG AAGAAGATAA GCCTCGATCC TTGGGTGACG AGTATCTATA 7800  
 TGTATATATG TACTATATGT TATGTTGTAC AGAAGAAGTC GTGTCCACAA ATATCAATTG 7860  
 ATGTCAGATG TCTAGTAAGT GATCATGTGT AGCATAAAA CTGGAGTAAT TAAAAAGTG 7920

AATAAACAAA AATAATTACT AAACGTTATT CCAAGTAGCT TTCCAAGACA GTCACTTGCC 7980  
 CTTTTCCAAT TTCCCTTGCA ATTAACATAA TTGCTCTTCA CGATATGATA TTATACCAAA 8040  
 ATGGTGATAC CTTGGGAATT GTTAATTTGA CTCATTTGAA CAAATCTCAT CTATAAAATC 8100  
 ATCCACCTC TCCACCACAT TTGTTCTCAC TACCAATCAA AAAATAATCT AGTCTTAAAC 8160

(2) INFORMATION FOR SEQ ID NO:111:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3195 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

ATGAATTCAA ATGGCCATGA GGAGGAGAAG AAGTTGGGGA ATGGAGTTGT GGGGATTCTA 60  
 TCTGAAACAG TTAACAAATG GGAGAGACGA ACACCATTGA CGCCATCGCA TTGCGCTCGC 120  
 CTTTTACACG GTGGGAAAGA CAGAACCGGC ATTTCCCGCA TTGTGGTTCA GCCATCTGCT 180  
 AAGCGTATCC ATCATGATGC CTTGTATGAA CATGTTGGGT GTGAAATTC TGATGATTTG 240  
 TCTGATTGTG GGCTTATACT TGAATCAAA CAACCTGAGC TAGAAATGAT TCTTCCAGAG 300  
 AGAGCATACG CTTTCTTTTC ACATACTCAT AAGGCACAGA AAGAGAACAT GCCTTTGTTG 360  
 GATAAAATTC TTTCTGAGAG AGTGACTIONG TGTGATTATG AGCTCATTGT TGGGGATCAT 420  
 GGGAAACGAT TATTGGCGTT TGGTAAATAT GCAGGCAGAG CTGGTCTTGT TGACTIONTTA 480  
 CACGGACTTG GACAGCGATA TCTAAGTCTA GGATACTCAA CACCTTTCCCT CTCGCTCGGT 540  
 GCATCGTATA TGTATTCTTC ATTTGGCTGCT GCAAAGCCG CTGTAATTTTC TGTGGTGAA 600  
 GAAATTGCAA GCCAGGGACT GCCATTAGGA ATCTGCCCTC TTGTATTTGT CTTACCGGA 660  
 ACAGGAAATG TTTCTCTGGG GGCACAAGAA ATTTTCAAGC TTCTTCCTCA CACTTTTGTT 720  
 GAACCAAGCA AACTTCCTGA ACTATTTGTA AAAGACAAAG GAATTAGTCA AAATGGGATT 780  
 TCAACAAAGC GAGTCTATCA AGTATATGGT TGTATTATTA CCAGCCAAGA CATGGTTGAA 840  
 CACAAAGATC CATCAAAGTC ATTCGACAAA GCCGACTATT ATGCACACCC GGAACATTAC 900  
 AATCCAGTTT TCCACGAAA GATATCGCCA TATACGTCTG TTCTTGTAAG CTGTATGTAC 960  
 TGGGAGAAGA GGTTCCTG TCTTCTGAGC ACAAACAGC TTCAAGATTT AACAAAAAAA 1020  
 GGACTCCCAC TAGTAGGCAT ATGTGATATA ACTTGTGACA TCGGTGGCTC CATTGAATTT 1080  
 GTTAACCGAG CTACTIONTAAAT CGATTCCCTT TTCTTCAGGT TTAATCCCTC GAACAATTCA 1140

TACTACGATG	ACATGGATGG	GGATGGCGTA	CTATGCATGG	CTGTTGACAT	TTTACCCACA	1200
GAATTTGCAA	AAGAGGCATC	CCAGCATTTT	GGAGATATTC	TTTCCGGATT	TGTCGGTAGT	1260
TTGGCTTCAA	TGACTGAAAT	TTCAGATCTA	CCAGCACATC	TGAAGAGGGC	TTGCATAAGC	1320
TATAGGGGAG	AATTGACATC	TTTGTATGAG	TATATTCCAC	GTATGAGGAA	GTCAAATCCA	1380
GAAGAGGCAC	AAGATAATAT	TATCGCCAAC	GGGGTTTCCA	GCCAGAGAAC	ATTCAACATA	1440
TTGGTATCTC	TGAGCGGACA	CCTATTTGAT	AAGTTTCTGA	TAAACGAAGC	TCTTGATATG	1500
ATCGAAGCGG	CTGGTGGCTC	ATTTCATTTG	GCTAAATGTG	AACTGGGGCA	GAGCGCTGAT	1560
GCTGAATCGT	ACTCAGAACT	TGAAGTTGGT	GCGGATGATA	AGAGAGTATT	GGATCAAATC	1620
ATTGATTCAT	TAACTCGGTT	AGCTAATCCA	AATGAAGATT	ATATATCCCC	ACATAGAGAA	1680
GCAAATAAGA	TCTCACTGAA	GATTGGTAAA	GTCCAGCAAG	AAAATGAGAT	AAAAGAGAAG	1740
CCTGAAATGA	CGAAAAATC	AGGTGTTTTG	ATTCTTGGTG	CTGGACGTGT	GTGTCGCCCA	1800
GCTGCTGATT	TCCTAGCTTC	AGTTAGAACC	ATTTCTGCAC	AGCAATGGTA	CAAAACATAT	1860
TTCCGGAGCAG	ACTCTGAAGA	GAAAACAGAT	GTTTCATGTGA	TTGTCGCGTC	TCTGTATCTT	1920
AAGGATGCCA	AAGAGACGGT	TGAAGGTATT	TCAGATGTAG	AAGCAGTTCG	GCTAGATGTA	1980
TCTGATAGTG	AAAGTCTCCT	TAAGTATGTT	TCTCAGGTTG	ATGTTGTCCT	AAGTTTATTA	2040
CCTGCAAGTT	GTCATGCTGT	TGTAGCAAAG	ACATGCATTG	AGCTGAAGAA	GCATCTCGTC	2100
ACTGCTAGCT	ATGTTGATGA	TGAAACGTCC	ATGTTACATG	AGAAGGCTAA	GAGTGCTGGG	2160
ATAACGATTC	TAGGCCAAAT	GGGACTGGAC	CCTGGAATCG	ATCACATGAT	GGCGATGAAA	2220
ATGATCAACG	ATGCTCATAT	CAAAAAGGG	AAAGTGAAGT	CTTTTACCTC	TTATTGTGGA	2280
GGGCTTCCCT	CTCCTGCTGC	AGCAAATAAT	CCATTAGCAT	ATAAATTTAG	CTGGAACCCT	2340
GCTGGAGCAA	TTCGAGCTGG	TCAAAACCCC	GCCAAATACA	AAAGCAACGG	CGACATAATA	2400
CATGTTGATG	GGAAGAATCT	CTATGATTCC	GCGGCAAGAT	TCCGAGTACC	TAATCTTCCA	2460
GCTTTTGCAT	TGGAGTGTTT	TCCAAATCGT	GACTCCTTGG	TTTACGGGGA	ACATTATGGC	2520
ATCGAGAGCG	AAGCAACAAC	GATATTTCTG	GGAACACTCA	GATATGAAGG	GTTTAGTATG	2580
ATAATGGCAA	CACTTTCGAA	ACTTGGATTC	TTTGACAGTG	AAGCAAATCA	AGTACTCTCC	2640
ACTGGAAAGA	GGATTACGTT	TGGTGCTCTT	TTAAGTAACA	TTCTAAATAA	GGATGCAGAC	2700
AATGAATCAG	AGCCCCTAGC	GGGAGAAGAA	GAGATAAGCA	AGAGAATTAT	CAAGCTTGGA	2760
CATTCCAAGG	AGACTGCAGC	CAAAGCTGCC	AAAACAATTG	TATTCTTGGG	GTTCAACGAA	2820
GAGAGGGAGG	TTCCATCACT	GTGTAAAAGC	GTATTTGATG	CAACTTGTTA	CCTAATGGAA	2880

GAGAAACTAG CTTATTCCGG AAATGAACAG GACATGGTGC TTTTGCATCA CGAAGTAGAA 2940  
 GTGGAATTCC TTGAAAGCAA ACGTATAGAG AAGCACACTG CGACTCTTTT GGAATTCCGGG 3000  
 GACATCAAGA ATGGACAAAC AACCAACCGCT ATGGCCAAGA CTGTTGGGAT CCCTGCAGCC 3060  
 ATTGGAGCTC TGGTGTTAAT TGAAGACAAG ATCAAGACAA GAGGAGTCTT AAGGCCTCTC 3120  
 GAAGCAGAGG TGTATTTGCC AGCTTTGGAT ATATTGCAAG CATATGGTAT AAAGCTGATG 3180  
 GAGAAGGCAG AATGA 3195

(2) INFORMATION FOR SEQ ID NO:112:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1064 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Met	Asn	Ser	Asn	Gly	His	Glu	Glu	Glu	Lys	Lys	Leu	Gly	Asn	Gly	Val	1	5	10	15
Val	Gly	Ile	Leu	Ser	Glu	Thr	Val	Asn	Lys	Trp	Glu	Arg	Arg	Thr	Pro	20	25	30	
Leu	Thr	Pro	Ser	His	Cys	Ala	Arg	Leu	Leu	His	Gly	Gly	Lys	Asp	Arg	35	40	45	
Thr	Gly	Ile	Ser	Arg	Ile	Val	Val	Gln	Pro	Ser	Ala	Lys	Arg	Ile	His	50	55	60	
His	Asp	Ala	Leu	Tyr	Glu	His	Val	Gly	Cys	Glu	Ile	Ser	Asp	Asp	Leu	65	70	75	80
Ser	Asp	Cys	Gly	Leu	Ile	Leu	Gly	Ile	Lys	Gln	Pro	Glu	Leu	Glu	Met	85	90	95	
Ile	Leu	Pro	Glu	Arg	Ala	Tyr	Ala	Phe	Phe	Ser	His	Thr	His	Lys	Ala	100	105	110	
Gln	Lys	Glu	Asn	Met	Pro	Leu	Leu	Asp	Lys	Ile	Leu	Ser	Glu	Arg	Val	115	120	125	
Thr	Leu	Cys	Asp	Tyr	Glu	Leu	Ile	Val	Gly	Asp	His	Gly	Lys	Arg	Leu	130	135	140	
Leu	Ala	Phe	Gly	Lys	Tyr	Ala	Gly	Arg	Ala	Gly	Leu	Val	Asp	Phe	Leu	145	150	155	160
His	Gly	Leu	Gly	Gln	Arg	Tyr	Leu	Ser	Leu	Gly	Tyr	Ser	Thr	Pro	Phe	165	170	175	



Leu Ser Leu Gly Ala Ser Tyr Met Tyr Ser Ser Leu Ala Ala Ala Lys  
 180 185 190

Ala Ala Val Ile Ser Val Gly Glu Glu Ile Ala Ser Gln Gly Leu Pro  
 195 200 205

Leu Gly Ile Cys Pro Leu Val Phe Val Phe Thr Gly Thr Gly Asn Val  
 210 215 220

Ser Leu Gly Ala Gln Glu Ile Phe Lys Leu Leu Pro His Thr Phe Val  
 225 230 235 240

Glu Pro Ser Lys Leu Pro Glu Leu Phe Val Lys Asp Lys Gly Ile Ser  
 245 250 255

Gln Asn Gly Ile Ser Thr Lys Arg Val Tyr Gln Val Tyr Gly Cys Ile  
 260 265 270

Ile Thr Ser Gln Asp Met Val Glu His Lys Asp Pro Ser Lys Ser Phe  
 275 280 285

Asp Lys Ala Asp Tyr Tyr Ala His Pro Glu His Tyr Asn Pro Val Phe  
 290 295 300

His Glu Lys Ile Ser Pro Tyr Thr Ser Val Leu Val Asn Cys Met Tyr  
 305 310 315 320

Trp Glu Lys Arg Phe Pro Cys Leu Leu Ser Thr Lys Gln Leu Gln Asp  
 325 330 335

Leu Thr Lys Lys Gly Leu Pro Leu Val Gly Ile Cys Asp Ile Thr Cys  
 340 345 350

Asp Ile Gly Gly Ser Ile Glu Phe Val Asn Arg Ala Thr Leu Ile Asp  
 355 360 365

Ser Pro Phe Phe Arg Phe Asn Pro Ser Asn Asn Ser Tyr Tyr Asp Asp  
 370 375 380

Met Asp Gly Asp Gly Val Leu Cys Met Ala Val Asp Ile Leu Pro Thr  
 385 390 395 400

Glu Phe Ala Lys Glu Ala Ser Gln His Phe Gly Asp Ile Leu Ser Gly  
 405 410 415

Phe Val Gly Ser Leu Ala Ser Met Thr Glu Ile Ser Asp Leu Pro Ala  
 420 425 430

His Leu Lys Arg Ala Cys Ile Ser Tyr Arg Gly Glu Leu Thr Ser Leu  
 435 440 445

Tyr Glu Tyr Ile Pro Arg Met Arg Lys Ser Asn Pro Glu Glu Ala Gln  
 450 455 460

Asp Asn Ile Ile Ala Asn Gly Val Ser Ser Gln Arg Thr Phe Asn Ile  
 465 470 475 480

Leu Val Ser Leu Ser Gly His Leu Phe Asp Lys Phe Leu Ile Asn Glu  
 485 490 495  
 Ala Leu Asp Met Ile Glu Ala Ala Gly Gly Ser Phe His Leu Ala Lys  
 500 505 510  
 Cys Glu Leu Gly Gln Ser Ala Asp Ala Glu Ser Tyr Ser Glu Leu Glu  
 515 520 525  
 Val Gly Ala Asp Asp Lys Arg Val Leu Asp Gln Ile Ile Asp Ser Leu  
 530 535 540  
 Thr Arg Leu Ala Asn Pro Asn Glu Asp Tyr Ile Ser Pro His Arg Glu  
 545 550 555 560  
 Ala Asn Lys Ile Ser Leu Lys Ile Gly Lys Val Gln Gln Glu Asn Glu  
 565 570 575  
 Ile Lys Glu Lys Pro Glu Met Thr Lys Lys Ser Gly Val Leu Ile Leu  
 580 585 590  
 Gly Ala Gly Arg Val Cys Arg Pro Ala Ala Asp Phe Leu Ala Ser Val  
 595 600 605  
 Arg Thr Ile Ser Ser Gln Gln Trp Tyr Lys Thr Tyr Phe Gly Ala Asp  
 610 615 620  
 Ser Glu Glu Lys Thr Asp Val His Val Ile Val Ala Ser Leu Tyr Leu  
 625 630 635 640  
 Lys Asp Ala Lys Glu Thr Val Glu Gly Ile Ser Asp Val Glu Ala Val  
 645 650 655  
 Arg Leu Asp Val Ser Asp Ser Glu Ser Leu Leu Lys Tyr Val Ser Gln  
 660 665 670  
 Val Asp Val Val Leu Ser Leu Leu Pro Ala Ser Cys His Ala Val Val  
 675 680 685  
 Ala Lys Thr Cys Ile Glu Leu Lys Lys His Leu Val Thr Ala Ser Tyr  
 690 695 700  
 Val Asp Asp Glu Thr Ser Met Leu His Glu Lys Ala Lys Ser Ala Gly  
 705 710 715 720  
 Ile Thr Ile Leu Gly Glu Met Gly Leu Asp Pro Gly Ile Asp His Met  
 725 730 735  
 Met Ala Met Lys Met Ile Asn Asp Ala His Ile Lys Lys Gly Lys Val  
 740 745 750  
 Lys Ser Phe Thr Ser Tyr Cys Gly Gly Leu Pro Ser Pro Ala Ala Ala  
 755 760 765  
 Asn Asn Pro Leu Ala Tyr Lys Phe Ser Trp Asn Pro Ala Gly Ala Ile  
 770 775 780

Arg Ala Gly Gln Asn Pro Ala Lys Tyr Lys Ser Asn Gly Asp Ile Ile  
 785 790 795 800  
 His Val Asp Gly Lys Asn Leu Tyr Asp Ser Ala Ala Arg Phe Arg Val  
 805 810 815  
 Pro Asn Leu Pro Ala Phe Ala Leu Glu Cys Phe Pro Asn Arg Asp Ser  
 820 825 830  
 Leu Val Tyr Gly Glu His Tyr Gly Ile Glu Ser Glu Ala Thr Thr Ile  
 835 840 845  
 Phe Arg Gly Thr Leu Arg Tyr Glu Gly Phe Ser Met Ile Met Ala Thr  
 850 855 860  
 Leu Ser Lys Leu Gly Phe Phe Asp Ser Glu Ala Asn Gln Val Leu Ser  
 865 870 875 880  
 Thr Gly Lys Arg Ile Thr Phe Gly Ala Leu Leu Ser Asn Ile Leu Asn  
 885 890 895  
 Lys Asp Ala Asp Asn Glu Ser Glu Pro Leu Ala Gly Glu Glu Glu Ile  
 900 905 910  
 Ser Lys Arg Ile Ile Lys Leu Gly His Ser Lys Glu Thr Ala Ala Lys  
 915 920 925  
 Ala Ala Lys Thr Ile Val Phe Leu Gly Phe Asn Glu Glu Arg Glu Val  
 930 935 940  
 Pro Ser Leu Cys Lys Ser Val Phe Asp Ala Thr Cys Tyr Leu Met Glu  
 945 950 955 960  
 Glu Lys Leu Ala Tyr Ser Gly Asn Glu Gln Asp Met Val Leu Leu His  
 965 970 975  
 His Glu Val Glu Val Glu Phe Leu Glu Ser Lys Arg Ile Glu Lys His  
 980 985 990  
 Thr Ala Thr Leu Leu Glu Phe Gly Asp Ile Lys Asn Gly Gln Thr Thr  
 995 1000 1005  
 Thr Ala Met Ala Lys Thr Val Gly Ile Pro Ala Ala Ile Gly Ala Leu  
 1010 1015 1020  
 Val Leu Ile Glu Asp Lys Ile Lys Thr Arg Gly Val Leu Arg Pro Leu  
 1025 1030 1035 1040  
 Glu Ala Glu Val Tyr Leu Pro Ala Leu Asp Ile Leu Gln Ala Tyr Gly  
 1045 1050 1055  
 Ile Lys Leu Met Glu Lys Ala Glu  
 1060

(2) INFORMATION FOR SEQ ID NO:113:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: modified\_base
  - (B) LOCATION: 6
  - (D) OTHER INFORMATION: /mod\_base=i
- (ix) FEATURE:
  - (A) NAME/KEY: modified\_base
  - (B) LOCATION: 12
  - (D) OTHER INFORMATION: /mod\_base=i
- (ix) FEATURE:
  - (A) NAME/KEY: modified\_base
  - (B) LOCATION: 21
  - (D) OTHER INFORMATION: /mod\_base=i
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

TTYTCNCAYA CNCAYAARGC NCA

23

(2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

TTYTCCART ACATRCARTT

20

(2) INFORMATION FOR SEQ ID NO:115:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 619 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

GAAAACATGC CTTTGCTGGA TAAGATTCTA GCTGAGAGGG CATCGTTATA TGA CTATGAA 60

TTAATTGTTG GGGACACTGG GAAAAGGTTA CTTGCATTTG GAAAATTCGC TGGTAGGGCT 120

GGAATGATCG ACTTTTTGCG CGGATTAGGA CAGCGGTTTT TAAGTCTTGG ATATTCAACA 180  
 CCTTCTTGT CACTTGGATC ATCTTACATG TACCCTTCCC TGGCTGCTGC TAAGGCTGCT 240  
 GTGATTTCTG TTGGTGAAAA ATTGCGACGC AGGGATTGCC ATTGGGGATT TGTCCCCTGG 300  
 TTTGTTTATT TACTGGTTCA GGAAATGTTT GTTCTGGTGC ACAGGAGATA TTTAAGCTTC 360  
 TTCTCATAAC CTTTGTGAT CCATCTAAAC TACGCGACCT ACATAGAACG GACCCAGATC 420  
 AACCAAGGCA TGCTTCAAAA AGAGTTTTCC AAGTTTATGG TTGTGTTGTG ACTGCCCAAG 480  
 ACATGGTTGA ACCCAAAGAT CACGTGATAG TGTTTGACAA AGCAGACTAC TATGCACATC 540  
 CTGAGCATT A CAATCCCCT TTCCATGAAA AAATAGCACC ATATGCATCT GTTATTGTCA 600  
 ATTGCATGTA TTGGGAAAA 619

(2) INFORMATION FOR SEQ ID NO:116:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 620 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GAGAATATGC CACTGTTAGA CAAGATCCTT GAAGAAAGGG TGCCTTGTG TGATTATGAG 60  
 CTAATTGTTG GAGATGATGG GAAAAGATCA CTAGCATTTG GGAAATTTGC TGGTAGAGCT 120  
 GGACTGATAG ATTTCTTACA TGGTCTCGGA CAGCGATATT TGAGCCTTGG ATACTCCACT 180  
 CCATTTCTCT CTCTGGGACA TCTCATATGT TCCTTCGCTC GCTGCAGCCA AGGCTGCAGT 240  
 CATTGTCGTT GCAGAAGAGA TAGCAACATT TGGACTTCCA TCCGGAATTT GTCCGATAGT 300  
 GTTTGTGTTT ACTGGAGTTG GAAACGTCTC TCAGGGTGCG CAGGAGATAT TCAAGTTATT 360  
 GCCCCATACC TTTGTTGATG CTGAGAAGCT TCCCAGAAAT TTTAGGCCA GGAATCTGTC 420  
 TAAGCAATCT CAGTCGACCA AGAGAGTATT TCAACTTTAT GGTGTTGTTG TGACCTCTAG 480  
 AGACATAGTT TCTCACAAGG ATCCCACCAG ACAATTTGAC AAAGGTGACT ATTATGCTCA 540  
 TCCAGAACAC TACACCCTG TTTTTCATGA AAGAATTGCT CCATATGCAT CTGTCATCGT 600  
 AACTGCATG TATTGGGAAA 620

(2) INFORMATION FOR SEQ ID NO:117:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 206 amino acids
  - (B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Glu Asn Met Pro Leu Leu Asp Lys Ile Leu Ala Glu Arg Ala Ser Leu  
1 5 10 15  
Tyr Asp Tyr Glu Leu Ile Val Gly Asp Thr Gly Lys Arg Leu Leu Ala  
20 25 30  
Phe Gly Lys Phe Ala Gly Arg Ala Gly Met Ile Asp Phe Leu Arg Gly  
35 40 45  
Leu Gly Gln Arg Phe Leu Ser Leu Gly Tyr Ser Thr Pro Phe Leu Ser  
50 55 60  
Leu Gly Ser Ser Tyr Met Tyr Pro Ser Leu Ala Ala Ala Lys Ala Ala  
65 70 75 80  
Val Ile Ser Val Gly Glu Xaa Ile Ala Thr Gln Gly Leu Pro Leu Gly  
85 90 95  
Ile Cys Pro Leu Val Cys Leu Phe Thr Gly Ser Gly Asn Val Cys Ser  
100 105 110  
Gly Ala Gln Glu Ile Phe Lys Leu Leu Pro His Thr Phe Val Asp Pro  
115 120 125  
Ser Lys Leu Arg Asp Leu His Arg Thr Asp Pro Asp Gln Pro Arg His  
130 135 140  
Ala Ser Lys Arg Val Phe Gln Val Tyr Gly Cys Val Val Thr Ala Gln  
145 150 155 160  
Asp Met Val Glu Pro Lys Asp His Val Ile Val Phe Asp Lys Ala Asp  
165 170 175  
Tyr Tyr Ala His Pro Glu His Tyr Asn Pro Thr Phe His Glu Lys Ile  
180 185 190  
Ala Pro Tyr Ala Ser Val Ile Val Asn Cys Met Tyr Trp Glu  
195 200 205

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 207 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

TTGAACCCAA AGATCACGTG ATAGTGTGTTG ACAAAGCAGA CTA CTATTCA CACCCTGAGC 60  
ATTACAATCC CACTTTCCAT GAAAAAATAG CACCATATGC ATCTGTTATT GTCAATTGCA 120  
TGTATTGGGA GAAAAGATTT CCTCAATTGC CGAGCTATAA GCAGATGCAA GACTTAATGG 180  
GCCGGGGGAG CCCCCTTGTT GGAATAGCTG ACATAACGTG TGATATAGGG GGTTC AATTG 240  
AGTTTGTTAA CCGCGGTACT TCAATTGATT CACCCTTCTT CAGATATGAT CCCTTAACAA 300  
ATTCC TACCA TGATGATATG GAGGGGAATG GAGTGATATG CTTAGCTGTT GACATTCTTC 360  
CAACAGAATT TGCAAAGGAG GCTTCCAAC ATTTTGAAA CATACTTCC CAATTTGTTG 420  
TAAATTTGGC TTCTGCTACA GACATTACAA AGTTGCCTGC TCACTTAAGG AGAGCTTGCA 480  
TAGCCATAA AGGAGTGCTA ACCTCCTTAT ATGATTATAT CCCACGCATG CGGAGTTCTG 540  
ATTCAGAGGA AGTATCAGAA AACGCAGAAA ATTCTCTATC CAACAAAAGG AAGTACAATA 600  
TATCGGTGTC TCTGAGTGGT CACTTATTTG ATCAGTTTCT GATAAATGAG GCCTTAGATA 660  
TTATTGAAGC TGCAGGAGGC TCCTTCCACT TAGTCAACTG CCATGTGGGT CAGAGCATTG 720  
AAGCCGTATC ATTCTCTGAA CTTGAAGTTG GTGCAGATAA CAGGGCTGTT CTGGATCAAA 780  
TCATTGATTC TTTAACTGCT ATTGCTAGTC CAACTGAACA TGATAGATTT TCAAATCAAG 840  
ATTCAAGTAA AATTTCACTT AAGCTTGGTA AAGTTGAAGA GAATGGCATA GAGAAGGAAT 900  
CTGACCCAG AAAGAAGGCT GCGGTTTTAA TTCTTGGAGC TGGTCGGGTC TGCAACCAG 960  
CTGCTGAAAT GTTATCATCA TTTGGAAGGC CATCATCGAG CCAATGGTAT AAAACATTGT 1020  
TGGAAGATGA TTTTGAATGT CAAACTGATG TAGAAGTCAT TGTGGGATCT CTGTACCTGA 1080  
AGGATGCAGA GCAGACTGTT GAGGGCATT CAAATGTAAC CGGAATTCAG CTTGATGTGA 1140  
TGGATCGTGC CAATTTGTGT AAGTACATTT CACAGGTTGA CGTTGTTATA AGTTTGCTGC 1200  
CCCCAAGTTG TCATATTATT GTAGCAAATG CTTGCATTGA GCTGAAAAAA CATCTTGCTCA 1260  
CTGCTAGCTA TGTTGATAGC TCCATGTCAA TGCTAAATGA TAAGGCTAAA GATGCTGGCA 1320  
TAACAATTCT TGGAGAGATG GGCTTGGACC CAGGAATTGG TCATATGATG GCAATGAAGA 1380  
TGATCAACCA AGCACATGTG AGGAAGGGGA AAATAAAGTC TTTCACTTCT TATTGTGGTG 1440  
GACTTCCATC TCCTGAAGCT GCTAACAATC CATTAGCATA TAAATTCAGT TGG AATCCTG 1500  
CAGGAGCCAT CCGAGCTGGG CGCAATCCTG CCACCTACAA ATGGGGTGGT GAAACTGTAC 1560  
ATATTGATGG GGACGATCTT TATGATTCGG CTACAAGACT AAGGCTACCG GACCTTCTG 1620  
CTTTTGCTTT GGAATGTCTC CCAAATCGCA ATTCATTA CT TATGGGGAT TTGTATGGAA 1680  
TAACTGAAGC ATCAACCATT TTCCGTGGAA CCCTCCGCTA TGAAGGATTT AGTGAGATCA 1740  
TGGGGACACT GTCTAGGATT AGCTTATTTA ACAATGAAGC CCATTCGTTG CTAATGAATG 1800  
GACAAAGACC AACTTTCAAA AAATTCCTAT TTGAACTTCT CAAAGTTGTT GGTGATAATC 1860



CAGATGAACT ATTGATAGGA GAGAATGACA TCATGGAGCA AATATTAATA CAAGGGCACT 1920  
 GCAAAGATCA AAGAACGGCA ATGGAGACAG CAAAAACAAT CATTTTCTTG GGACTTCTTG 1980  
 ACCAAACTGA AATCCCTGCT TCCTGCAAAA GTGCTTTTGA TGTTGCTTGT TTCCGCATGG 2040  
 AGGAGAGGTT ATCATAACACC AGCACAGAAA AGGATATGGT GCTTTTGCAT CATGAAGTGG 2100  
 AAATAGAATA CCCAGATAGC CAAATTACAG AGAAGCATAG AGCTACTTTA CTTGAATTTG 2160  
 GGAAGACTCT TGATGAAAAA ACCACAACCTG CCATGGCCCT TACTGTTGGT ATTCCAGCTG 2220  
 CTGTTGGAGC TTTGCTTTTA TTGACAAACA AAATTCAGAC AAGAGGAGTC TTAAGGCCTA 2280  
 TCGAACCTGA AGTATACAAT CCAGCACTGG ATATTATAGA AGCTTATGGG ATCAAGTTGA 2340  
 TAGAGAAGAC CGAGTAATTT GCATYTATGA ATTGATGTAT AGGTGTACAT TAATGTACAC 2400  
 CATGCAATGT TTGATTTGAA TAAGATAAAA TATAATAATT ACTGCAGTCA TGGAAATTGCA 2460  
 ACTGCCATTC TATGCAACTG TCAGAAATGG ACCACACGGT ACCAGCATAG TTA AAAACT 2520  
 TAGGCAGATA CCAATTTCAA TTGCAGCAGT ACAATCCAAC CAGTTATGAA GTATGGTTCT 2580  
 AG 2582

(2) INFORMATION FOR SEQ ID NO:120:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3265 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Zea mays
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 3..3071
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

ATTGTGCCCG CCTTCTGCTA GGAGGAGGCA AGAACGGACC TCGAGTAAAC CGGATTATTG 60  
 TGCAGCCAAG CACAAGGAGG ATCCATCATG ACGCTCAGTA TGAGGATGCA GGATGCGAGA 120  
 TTTCAGAAGA CCTGTCAGAA TGCGGCCTTA TCATAGGCAT CAAACAACCC AAGCTGCAGA 180  
 TGATTCTTTC AGATAGAGCG TACGCTTTCT TTTCACACAC ACACAAAGCC CAAAAGAGA 240  
 ATATGCCACT GTTAGACAAG ATCCTTGAAG AAAGGGTGTC CTTGTTTGAT TATGAGCTAA 300  
 TTGTTGGAGA TGATGGGAAA AGATCACTAG CATTTGGGAA ATTTGCTGGT AGAGCTGGAC 360  
 TGATAGATTT CTTACATGGT CTCGGACAGC GATATTTGAG CTTGGATAC TCGACTCCAT 420  
 TTCTCTCTCT GGGACAATCT CATATGTATC CTTCGCTCGC TGCAGCCAAG GCTGCAGTCA 480

TTGTCGTTGC	AGAAGAGATA	GCAACATTTG	GACTTCCATC	CGGAATTTGT	CCGATAGTGT	540
TTGTGTTTAC	TGGAGTTGGA	AACGTCTCTC	AGGGTGCGCA	GGAGATATTC	AAGTTATTGC	600
CCCATACCTT	TGTTGATGCT	GAGAAGCTTC	CCGAAATTTT	TCAGGCCAGG	AATCTGTCTA	660
AGCAATCTCA	GTCGACCAAG	AGAGTATTTT	AACTTTATGG	TTGTGTTGTG	ACCTCTAGAG	720
ACATAGTTTC	TCACAAGGAT	CCCACCAGAC	AATTTGACAA	AGGTGACTAT	TATGCTCATC	780
CAGAACACTA	CACCCCTGTT	TTTCATGAAA	GAATTGCTCC	ATATGCATCT	GTCATCGTAA	840
ACTGTATGTA	TTGGGAGAAG	AGGTTTCCAC	CATTACTAAA	TATGGATCAG	TTACAGCAAT	900
TGATGGAGAC	TGGTTGTCCT	TTAGTCGGCG	TTTGTGACAT	AACTTGTGAT	ATTGGAGGTT	960
CCATTGAATT	TATCAACAAG	AGTACATCAA	TAGAGAGGCC	TTTCTTTCGG	TATGATCCTT	1020
CTAAGAATTC	ATACCATGAT	GATATGGAAG	GTGCCGGAGT	GGTCTGCTTG	GCTGTTGACA	1080
TTCTCCCTAC	AGAATTCTCT	AAAGAGGCCT	CCCAACATTT	TGGAAACATA	CTATCTAGAC	1140
TTGTTGCTAG	TTTGGCCTCA	GTGAAGCAAC	CGGCAGAACT	TCCTTCCTAC	TTGAGAAGAG	1200
CTTGCATTGC	ACATGCTGGC	AGATTAACTC	CTTTGTATGA	ATATATCCCT	AGGATGAGAA	1260
ATACTATGAT	AGATTTGGCA	CCCGCAAAAA	CAAATCCATT	GCCTGACAAG	AAGTATAGCA	1320
CCCTGGTATC	TCTCAGTGGG	CACCTATTTG	ATAAGTTCCCT	TATAAATGAA	GCTTTGGACA	1380
TCATTGAGAC	AGCTGGAGGT	TCATTTCACT	TGGTTAGATG	TGAAGTTGGA	CAAAGCACGG	1440
ATGATATGTC	ATACTCAGAG	CTTGAAGTAG	GAGCAGATGA	TACTGCCACA	TTGGATAAAA	1500
TTATTGATTC	CTTGACTTCT	TTAGCTAATG	AACATGGTGG	AGATCACGAT	GCCGGGCAAG	1560
AAATTGAATT	AGCTCTGAAG	ATAGGAAAAG	TCAATGAGTA	TGAAACTGAC	GTCACAATTG	1620
ATAAAGGAGG	GCCAAAGATT	TTAATTCTTG	GAGCTGGAAG	AGTCTGTCCG	CCAGCTGCTG	1680
AGTTTCTGGC	ATCTTACCCA	GACATATGTA	CCTATGGTGT	TGATGACCAT	GATGCAGATC	1740
AAATTCATGT	TATCGTGGCA	TCTTTGTATC	AAAAAGATGC	AGAAGAGACA	GTTGATGGTA	1800
TTGAAAATAC	AACTGCTACC	CAGCTTGATG	TTGCTGATAT	TGGAAGCCTT	TCAGATCTTG	1860
TTTCTCAGGT	TGAGGTTGTA	ATTAGCTTGC	TGCCTGCTAG	TTTTCATGCT	GCCATTGCAG	1920
GAGTATGCAT	AGAGTTGAAG	AAGCACATGG	TAACGGCAAG	CTATGTTGAT	GAATCCATGT	1980
CAAACCTGAG	CCAAGCTGCC	AAAGATGCAG	GTGTAACTAT	ACTTTGTGAA	ATGGGCCTAG	2040
ATCCTGGCAT	AGATCACTTG	ATGTCAATGA	AGATGATTGA	TGAAGCTCAT	GCACGAAAGG	2100
GAAAAATAAA	GGCATTTACA	TCTTACTGTG	GTGGATTGCC	ATCTCCAGCT	GCAGCAAACA	2160
ATCCGCTTGC	CTATAAATTC	AGTTGGAACC	CAGCTGGTGC	ACTCCGGTCA	GGGAAAAATC	2220
CTGCAGTCTA	CAAATTTCTT	GGTGAGACGA	TCCATGTAGA	TGGTCATAAC	TTGTATGAAT	2280
CAGCAAAGAG	GCTCAGACTA	CGAGAGCTTC	CAGCTTTTGC	TCTGGAACAC	TTGCCAAATC	2340
GGAATTCCTT	GATATATGGT	GACCTTTATG	GSTATCTCAA	AGAAGCATCC	ACCATATATA	2400

GGGCTACTYT TCGTTACGAA GGTTTTAGTG AGATTATGGT AACCCTTTCC AAAACTGGGT 2460  
TCTTTGATGC TGCAAATCAT CCACTGCTGC AAGATACTAG TCGTCCAACA TATAAGGGTT 2520  
TCCTTGATGA ACTACTGAAT AATATCTCCA CAATTAACAC GGACTTAGAT ATTGAAGCTT 2580  
CTGGTGGATA CGATGATGAC CTGATTGCCA GACTGTTGAA GCTCGGGTGT TGCAAAAATA 2640  
AGGAAATAGC TGTTAAGACA GTCAAAACCA TCAAGTTCTT GGGACTACAT GAAGAGACTC 2700  
AAATACCTAA GGGTTGTTTCG AGCCCATTTG ATGTGATTTG CCAGCGAATG GAACAGAGGA 2760  
TGGCCTATGG CCACAATGAG CAAGACATGG TACTGCTCCA CCACGAAGTC GAGGTGGAAT 2820  
ACCCGGACGG GCAACCCGCC GAAAAGCACC AAGCGACGCT ACTGGAGTTC GGAAGGTTG 2880  
AAAATGGCAG GTCCACCACT GCCATGGCGC TGACCGTCGG CATTCAGCA GCAATAGGGG 2940  
CCCTGCTATT GCTAAAGAAT AAGGTCCAGA CGAAAGGAGT GATCAGGCCT CTGCAACCGG 3000  
AAATCTACGT TCCAGCATTG GAGATCTTGG AGTCGTCGGG CATCAAGCTG GTTGAGAAAG 3060  
TGGAGACTTG AAAGTTCCTT GATACACAGA TAAAGATAGT ATGATATAGC AGGGCACATG 3120  
TATCTTTTGT ATTAACCTCCG TTCTGGAATA TATATTTGTG AACTAAAATG TGACAAATAA 3180  
AAAGAACGGG TGGAGTATAT TGTAAGAGAC GGCAAAGAAA CCTCTGTATA TATGACCTGT 3240  
CGATATCAAA TAATGCCGAT CAGTT 3265

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 784 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Glycine max

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Glu Pro Lys Asp His Val Ile Val Phe Asp Lys Ala Asp Tyr Tyr Ser  
1 5 10 15  
His Pro Glu His Tyr Asn Pro Thr Phe His Glu Lys Ile Ala Pro Tyr  
20 25 30  
Ala Ser Val Ile Val Asn Cys Met Tyr Trp Glu Lys Arg Phe Pro Gln  
35 40 45  
Leu Pro Ser Tyr Lys Gln Met Gln Asp Leu Met Gly Arg Gly Ser Pro  
50 55 60  
Leu Val Gly Ile Ala Asp Ile Thr Cys Asp Ile Gly Gly Ser Ile Glu  
65 70 75 80

Phe Val Asn Arg Gly Thr Ser Ile Asp Ser Pro Phe Phe Arg Tyr Asp  
85 90 95  
Pro Leu Thr Asn Ser Tyr His Asp Asp Met Glu Gly Asn Gly Val Ile  
100 105 110  
Cys Leu Ala Val Asp Ile Leu Pro Thr Glu Phe Ala Lys Glu Ala Ser  
115 120 125  
Gln His Phe Gly Asn Ile Leu Ser Gln Phe Val Val Asn Leu Ala Ser  
130 135 140  
Ala Thr Asp Ile Thr Lys Leu Pro Ala His Leu Arg Arg Ala Cys Ile  
145 150 155 160  
Ala His Lys Gly Val Leu Thr Ser Leu Tyr Asp Tyr Ile Pro Arg Met  
165 170 175  
Arg Ser Ser Asp Ser Glu Glu Val Ser Glu Asn Ala Glu Asn Ser Leu  
180 185 190  
Ser Asn Lys Arg Lys Tyr Asn Ile Ser Val Ser Leu Ser Gly His Leu  
195 200 205  
Phe Asp Gln Phe Leu Ile Asn Glu Ala Leu Asp Ile Ile Glu Ala Ala  
210 215 220  
Gly Gly Ser Phe His Leu Val Asn Cys His Val Gly Gln Ser Ile Glu  
225 230 235 240  
Ala Val Ser Phe Ser Glu Leu Glu Val Gly Ala Asp Asn Arg Ala Val  
245 250 255  
Leu Asp Gln Ile Ile Asp Ser Leu Thr Ala Ile Ala Ser Pro Thr Glu  
260 265 270  
His Asp Arg Phe Ser Asn Gln Asp Ser Ser Lys Ile Ser Leu Lys Leu  
275 280 285  
Gly Lys Val Glu Glu Asn Gly Ile Glu Lys Glu Ser Asp Pro Arg Lys  
290 295 300  
Lys Ala Ala Val Leu Ile Leu Gly Ala Gly Arg Val Cys Gln Pro Ala  
305 310 315 320  
Ala Glu Met Leu Ser Ser Phe Gly Arg Pro Ser Ser Ser Gln Trp Tyr  
325 330 335  
Lys Thr Leu Leu Glu Asp Asp Phe Glu Cys Gln Thr Asp Val Glu Val  
340 345 350  
Ile Val Gly Ser Leu Tyr Leu Lys Asp Ala Glu Gln Thr Val Glu Gly  
355 360 365  
Ile Pro Asn Val Thr Gly Ile Gln Leu Asp Val Met Asp Arg Ala Asn  
370 375 380  
Leu Cys Lys Tyr Ile Ser Gln Val Asp Val Val Ile Ser Leu Leu Pro  
385 390 395 400  
Pro Ser Cys His Ile Ile Val Ala Asn Ala Cys Ile Glu Leu Lys Lys  
405 410 415  
His Leu Val Thr Ala Ser Tyr Val Asp Ser Ser Met Ser Met Leu Asn  
420 425 430

Asp Lys Ala Lys Asp Ala Gly Ile Thr Ile Leu Gly Glu Met Gly Leu  
 435 440 445  
 Asp Pro Gly Ile Gly His Met Met Ala Met Lys Met Ile Asn Gln Ala  
 450 455 460  
 His Val Arg Lys Gly Lys Ile Lys Ser Phe Thr Ser Tyr Cys Gly Gly  
 465 470 475 480  
 Leu Pro Ser Pro Glu Ala Ala Asn Asn Pro Leu Ala Tyr Lys Phe Ser  
 485 490 495  
 Trp Asn Pro Ala Gly Ala Ile Arg Ala Gly Arg Asn Pro Ala Thr Tyr  
 500 505 510  
 Lys Trp Gly Gly Glu Thr Val His Ile Asp Gly Asp Asp Leu Tyr Asp  
 515 520 525  
 Ser Ala Thr Arg Leu Arg Leu Pro Asp Leu Pro Ala Phe Ala Leu Glu  
 530 535 540  
 Cys Leu Pro Asn Arg Asn Ser Leu Leu Tyr Gly Asp Leu Tyr Gly Ile  
 545 550 555 560  
 Thr Glu Ala Ser Thr Ile Phe Arg Gly Thr Leu Arg Tyr Glu Gly Phe  
 565 570 575  
 Ser Glu Ile Met Gly Thr Leu Ser Arg Ile Ser Leu Phe Asn Asn Glu  
 580 585 590  
 Ala His Ser Leu Leu Met Asn Gly Gln Arg Pro Thr Phe Lys Lys Phe  
 595 600 605  
 Leu Phe Glu Leu Leu Lys Val Val Gly Asp Asn Pro Asp Glu Leu Leu  
 610 615 620  
 Ile Gly Glu Asn Asp Ile Met Glu Gln Ile Leu Ile Gln Gly His Cys  
 625 630 635 640  
 Lys Asp Gln Arg Thr Ala Met Glu Thr Ala Lys Thr Ile Ile Phe Leu  
 645 650 655  
 Gly Leu Leu Asp Gln Thr Glu Ile Pro Ala Ser Cys Lys Ser Ala Phe  
 660 665 670  
 Asp Val Ala Cys Phe Arg Met Glu Glu Arg Leu Ser Tyr Thr Ser Thr  
 675 680 685  
 Glu Lys Asp Met Val Leu Leu His His Glu Val Glu Ile Glu Tyr Pro  
 690 695 700  
 Asp Ser Gln Ile Thr Glu Lys His Arg Ala Thr Leu Leu Glu Phe Gly  
 705 710 715 720  
 Lys Thr Leu Asp Glu Lys Thr Thr Thr Ala Met Ala Leu Thr Val Gly  
 725 730 735  
 Ile Pro Ala Ala Val Gly Ala Leu Leu Leu Leu Thr Asn Lys Ile Gln  
 740 745 750  
 Thr Arg Gly Val Leu Arg Pro Ile Glu Pro Glu Val Tyr Asn Pro Ala  
 755 760 765  
 Leu Asp Ile Ile Glu Ala Tyr Gly Ile Lys Leu Ile Glu Lys Thr Glu

## (2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1022 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Zea mays

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Cys	Ala	Arg	Leu	Leu	Leu	Gly	Gly	Gly	Lys	Asn	Gly	Pro	Arg	Val	Asn
1			5						10					15	
Arg	Ile	Ile	Val	Gln	Pro	Ser	Thr	Arg	Arg	Ile	His	His	Asp	Ala	Gln
			20					25					30		
Tyr	Glu	Asp	Ala	Gly	Cys	Glu	Ile	Ser	Glu	Asp	Leu	Ser	Glu	Cys	Gly
		35					40					45			
Leu	Ile	Ile	Gly	Ile	Lys	Gln	Pro	Lys	Leu	Gln	Met	Ile	Leu	Ser	Asp
	50					55					60				
Arg	Ala	Tyr	Ala	Phe	Phe	Ser	His	Thr	His	Lys	Ala	Gln	Lys	Glu	Asn
	65				70					75					80
Met	Pro	Leu	Leu	Asp	Lys	Ile	Leu	Glu	Glu	Arg	Val	Ser	Leu	Phe	Asp
				85					90					95	
Tyr	Glu	Leu	Ile	Val	Gly	Asp	Asp	Gly	Lys	Arg	Ser	Leu	Ala	Phe	Gly
			100					105					110		
Lys	Phe	Ala	Gly	Arg	Ala	Gly	Leu	Ile	Asp	Phe	Leu	His	Gly	Leu	Gly
		115					120					125			
Gln	Arg	Tyr	Leu	Ser	Leu	Gly	Tyr	Ser	Thr	Pro	Phe	Leu	Ser	Leu	Gly
	130					135					140				
Gln	Ser	His	Met	Tyr	Pro	Ser	Leu	Ala	Ala	Ala	Lys	Ala	Ala	Val	Ile
	145				150					155					160
Val	Val	Ala	Glu	Glu	Ile	Ala	Thr	Phe	Gly	Leu	Pro	Ser	Gly	Ile	Cys
				165					170					175	
Pro	Ile	Val	Phe	Val	Phe	Thr	Gly	Val	Gly	Asn	Val	Ser	Gln	Gly	Ala
			180					185					190		
Gln	Glu	Ile	Phe	Lys	Leu	Leu	Pro	His	Thr	Phe	Val	Asp	Ala	Glu	Lys
		195					200					205			
Leu	Pro	Glu	Ile	Phe	Gln	Ala	Arg	Asn	Leu	Ser	Lys	Gln	Ser	Gln	Ser
	210					215					220				
Thr	Lys	Arg	Val	Phe	Gln	Leu	Tyr	Gly	Cys	Val	Val	Thr	Ser	Arg	Asp
	225				230					235					240
Ile	Val	Ser	His	Lys	Asp	Pro	Thr	Arg	Gln	Phe	Asp	Lys	Gly	Asp	Tyr

245							250					255			
Tyr	Ala	His	Pro	Glu	His	Tyr	Thr	Pro	Val	Phe	His	Glu	Arg	Ile	Ala
			260					265					270		
Pro	Tyr	Ala	Ser	Val	Ile	Val	Asn	Cys	Met	Tyr	Trp	Glu	Lys	Arg	Phe
		275					280					285			
Pro	Pro	Leu	Leu	Asn	Met	Asp	Gln	Leu	Gln	Gln	Leu	Met	Glu	Thr	Gly
	290					295						300			
Cys	Pro	Leu	Val	Gly	Val	Cys	Asp	Ile	Thr	Cys	Asp	Ile	Gly	Gly	Ser
305					310					315					320
Ile	Glu	Phe	Ile	Asn	Lys	Ser	Thr	Ser	Ile	Glu	Arg	Pro	Phe	Phe	Arg
				325					330						335
Tyr	Asp	Pro	Ser	Lys	Asn	Ser	Tyr	His	Asp	Asp	Met	Glu	Gly	Ala	Gly
			340					345					350		
Val	Val	Cys	Leu	Ala	Val	Asp	Ile	Leu	Pro	Thr	Glu	Phe	Ser	Lys	Glu
		355					360					365			
Ala	Ser	Gln	His	Phe	Gly	Asn	Ile	Leu	Ser	Arg	Leu	Val	Ala	Ser	Leu
	370					375					380				
Ala	Ser	Val	Lys	Gln	Pro	Ala	Glu	Leu	Pro	Ser	Tyr	Leu	Arg	Arg	Ala
385					390					395					400
Cys	Ile	Ala	His	Ala	Gly	Arg	Leu	Thr	Pro	Leu	Tyr	Glu	Tyr	Ile	Pro
				405					410					415	
Arg	Met	Arg	Asn	Thr	Met	Ile	Asp	Leu	Ala	Pro	Ala	Lys	Thr	Asn	Pro
			420					425					430		
Leu	Pro	Asp	Lys	Lys	Tyr	Ser	Thr	Leu	Val	Ser	Leu	Ser	Gly	His	Leu
		435					440					445			
Phe	Asp	Lys	Phe	Leu	Ile	Asn	Glu	Ala	Leu	Asp	Ile	Ile	Glu	Thr	Ala
	450					455					460				
Gly	Gly	Ser	Phe	His	Leu	Val	Arg	Cys	Glu	Val	Gly	Gln	Ser	Thr	Asp
465					470					475					480
Asp	Met	Ser	Tyr	Ser	Glu	Leu	Glu	Val	Gly	Ala	Asp	Asp	Thr	Ala	Thr
				485					490					495	
Leu	Asp	Lys	Ile	Ile	Asp	Ser	Leu	Thr	Ser	Leu	Ala	Asn	Glu	His	Gly
			500					505					510		
Gly	Asp	His	Asp	Ala	Gly	Gln	Glu	Ile	Glu	Leu	Ala	Leu	Lys	Ile	Gly
		515					520					525			
Lys	Val	Asn	Glu	Tyr	Glu	Thr	Asp	Val	Thr	Ile	Asp	Lys	Gly	Gly	Pro
	530					535					540				
Lys	Ile	Leu	Ile	Leu	Gly	Ala	Gly	Arg	Val	Cys	Arg	Pro	Ala	Ala	Glu
545					550					555					560
Phe	Leu	Ala	Ser	Tyr	Pro	Asp	Ile	Cys	Thr	Tyr	Gly	Val	Asp	Asp	His
				565					570					575	
Asp	Ala	Asp	Gln	Ile	His	Val	Ile	Val	Ala	Ser	Leu	Tyr	Gln	Lys	Asp
			580					585					590		

Ala Glu Glu Thr Val Asp Gly Ile Glu Asn Thr Thr Ala Thr Gln Leu  
595 600 605  
Asp Val Ala Asp Ile Gly Ser Leu Ser Asp Leu Val Ser Gln Val Glu  
610 615 620  
Val Val Ile Ser Leu Leu Pro Ala Ser Phe His Ala Ala Ile Ala Gly  
625 630 635 640  
Val Cys Ile Glu Leu Lys Lys His Met Val Thr Ala Ser Tyr Val Asp  
645 650 655  
Glu Ser Met Ser Asn Leu Ser Gln Ala Ala Lys Asp Ala Gly Val Thr  
660 665 670  
Ile Leu Cys Glu Met Gly Leu Asp Pro Gly Ile Asp His Leu Met Ser  
675 680 685  
Met Lys Met Ile Asp Glu Ala His Ala Arg Lys Gly Lys Ile Lys Ala  
690 695 700  
Phe Thr Ser Tyr Cys Gly Gly Leu Pro Ser Pro Ala Ala Ala Asn Asn  
705 710 715 720  
Pro Leu Ala Tyr Lys Phe Ser Trp Asn Pro Ala Gly Ala Leu Arg Ser  
725 730 735  
Gly Lys Asn Pro Ala Val Tyr Lys Phe Leu Gly Glu Thr Ile His Val  
740 745 750  
Asp Gly His Asn Leu Tyr Glu Ser Ala Lys Arg Leu Arg Leu Arg Glu  
755 760 765  
Leu Pro Ala Phe Ala Leu Glu His Leu Pro Asn Arg Asn Ser Leu Ile  
770 775 780  
Tyr Gly Asp Leu Tyr Gly Ile Ser Lys Glu Ala Ser Thr Ile Tyr Arg  
785 790 795 800  
Ala Thr Xaa Arg Tyr Glu Gly Phe Ser Glu Ile Met Val Thr Leu Ser  
805 810 815  
Lys Thr Gly Phe Phe Asp Ala Ala Asn His Pro Leu Leu Gln Asp Thr  
820 825 830  
Ser Arg Pro Thr Tyr Lys Gly Phe Leu Asp Glu Leu Leu Asn Asn Ile  
835 840 845  
Ser Thr Ile Asn Thr Asp Leu Asp Ile Glu Ala Ser Gly Gly Tyr Asp  
850 855 860  
Asp Asp Leu Ile Ala Arg Leu Leu Lys Leu Gly Cys Cys Lys Asn Lys  
865 870 875 880  
Glu Ile Ala Val Lys Thr Val Lys Thr Ile Lys Phe Leu Gly Leu His  
885 890 895  
Glu Glu Thr Gln Ile Pro Lys Gly Cys Ser Ser Pro Phe Asp Val Ile  
900 905 910  
Cys Gln Arg Met Glu Gln Arg Met Ala Tyr Gly His Asn Glu Gln Asp  
915 920 925  
Met Val Leu Leu His His Glu Val Glu Val Glu Tyr Pro Asp Gly Gln  
930 935 940



Pro Ala Glu Lys His Gln Ala Thr Leu Leu Glu Phe Gly Lys Val Glu  
945 950 955 960

Asn Gly Arg Ser Thr Thr Ala Met Ala Leu Thr Val Gly Ile Pro Ala  
965 970 975

Ala Ile Gly Ala Leu Leu Leu Leu Lys Asn Lys Val Gln Thr Lys Gly  
980 985 990

Val Ile Arg Pro Leu Gln Pro Glu Ile Tyr Val Pro Ala Leu Glu Ile  
995 1000 1005

Leu Glu Ser Ser Gly Ile Lys Leu Val Glu Lys Val Glu Thr  
1010 1015 1020

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1908 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Zea mays

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 3..1908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

ATTGTGCCCG CCTTCTGCTA GGAGGAGGCA AGAACGGACC TCGAGTAAAC CGGATTATTG 60  
TGCAGCCAAG CACAAGGAGG ATCCATCATG ACGCTCAGTA TGAGGATGCA GGATGCGAGA 120  
TTTCAGAAGA CCTGTCAGAA TCGGGCCTTA TCATAGGCAT CAAACAACCC AAGCTGCAGA 180  
TGATTCTTTC AGATAGAGCG TACGCTTCT TTTACACAC ACACAAAGCC CAAAAGAGA 240  
ATATGCCACT GTTAGACAAG ATCCTTGAAG AAAGGGTGTC CTTGTTTGAT TATGAGCTAA 300  
TTGTTGGAGA TGATGGGAAA AGATCACTAG CATTGGGAA ATTTGCTGGT AGAGCTGGAC 360  
TGATAGATTT CTTACATGGT CTCGGACAGC GATATTTGAG CTTGGATAC TCGACTCCAT 420  
TTCTCTCTCT GGGACAATCT CATATGTATC CTTCGCTCGC TGCAGCCAAG GCTGCAGTCA 480  
TTGTCGTTGC AGAAGAGATA GCAACATTTG GACTTCCATC CGGAATTTGT CCGATAGTGT 540  
TTGTGTTTAC TGGAGTTGGA AACGTCTCTC AGGGTGCGCA GGAGATATTC AAGTTATTGC 600  
CCCATACCTT TGTTGATGCT GAGAAGCTTC CCGAAATTTT TCAGGCCAGG AATCTGTCTA 660  
AGCAATCTCA GTCGACCAAG AGAGTATTTT AACTTTATGG TTGTGTTGTG ACCTCTAGAG 720  
ACATAGTTTC TCACAAGGAT CCCACCAGAC AATTTGACAA AGGTGACTAT TATGCTCATC 780

CAGAACTA CACCCCTGTT TTTCATGAAA GAATTGCTCC ATATGCATCT GTCATCGTAA 840  
 ACTGTATGTA TTGGGAGAAG AGGTTTCCAC CATTACTAAA TATGGATCAG TTACAGCAAT 900  
 TGATGGAGAC TGGTTGTCCT TTAGTCGGCG TTTGTGACAT AACTTGTGAT ATTGGAGGTT 960  
 CCATTGAATT TATCAACAAG AGTACATCAA TAGAGAGGCC TTTCTTTTCGG TATGATCCTT 1020  
 CTAAGAATTC ATACCATGAT GATATGGAAG GTGCCGGAGT GGTCTGCTTG GCTGTTGACA 1080  
 TTCTCCCTAC AGAATTCTCT AAAGAGGCCT CCCAACATTT TGAAACATA CTATCTAGAC 1140  
 TTGTTGCTAG TTTGGCCTCA GTGAAGCAAC CGGCAGAACT TCCTTCCTAC TTGAGAAGAG 1200  
 CTTGCATTGC ACATGCTGGC AGATTAATC CTTTGTATGA ATATATCCCT AGGATGAGAA 1260  
 ATACTATGAT AGATTTGGCA CCCGCAAAAA CAAATCCATT GCCTGACAAG AAGTATAGCA 1320  
 CCCTGGTATC TCTCAGTGGG CACCTATTTG ATAAGTTCCT TATAAATGAA GCTTTGGACA 1380  
 TCATTGAGAC AGCTGGAGGT TCATTTCACT TGGTTAGATG TGAAGTTGGA CAAAGCACGG 1440  
 ATGATATGTC ATACTCAGAG CTTGAAGTAG GAGCAGATGA TACTGCCACA TTGGATAAAA 1500  
 TTATTGATTC CTTGACTTCT TTAGCTAATG AACATGGTGG AGATCACGAT GCCGGGCAAG 1560  
 AAATTGAATT AGCTCTGAAG ATAGGAAAAG TCAATGAGTA TGAAACTGAC GTCACAATTG 1620  
 ATAAAGGAGG GCCAAAGATT TTAATTCTTG GAGCTGGAAG AGTCTGTCCG CCAGCTGCTG 1680  
 AGTTTCTGGC ATCTTACCCA GACATATGTA CCTATGGTGT TGATGACCAT GATGCAGATC 1740  
 AAATTCATGT TATCGTGGCA TCTTTGTATC AAAAAGATGC AGAAGAGACA GTTGATGGTA 1800  
 TTGAAAATAC AACTGCTACC CAGCTTGATG TTGCTGATAT TGGAAGCCTT TCAGATCTTG 1860  
 TTTCTCAGGT TGAGGTTGTA ATTAGCTTGC TGCCTGCTAG TTTTCATG 1908

(2) INFORMATION FOR SEQ ID NO:124:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 640 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 

- (A) ORGANISM: Zea mays

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Cys Ala Arg Leu Leu Leu Gly Gly Gly Lys Asn Gly Pro Arg Val Asn  
 1                   5                   10                   15  
 Arg Ile Ile Val Gln Pro Ser Thr Arg Arg Ile His His Asp Ala Gln  
                  20                   25                   30

Tyr Glu Asp Ala Gly Cys Glu Ile Ser Glu Asp Leu Ser Glu Cys Gly  
 35 40 45  
 Leu Ile Ile Gly Ile Lys Gln Pro Lys Leu Gln Met Ile Leu Ser Asp  
 50 55 60  
 Arg Ala Tyr Ala Phe Phe Ser His Thr His Lys Ala Gln Lys Glu Asn  
 65 70 75 80  
 Met Pro Leu Leu Asp Lys Ile Leu Glu Glu Arg Val Ser Leu Phe Asp  
 85 90 95  
 Tyr Glu Leu Ile Val Gly Asp Asp Gly Lys Arg Ser Leu Ala Phe Gly  
 100 105 110  
 Lys Phe Ala Gly Arg Ala Gly Leu Ile Asp Phe Leu His Gly Leu Gly  
 115 120 125  
 Gln Arg Tyr Leu Ser Leu Gly Tyr Ser Thr Pro Phe Leu Ser Leu Gly  
 130 135 140  
 Gln Ser His Met Tyr Pro Ser Leu Ala Ala Ala Lys Ala Ala Val Ile  
 145 150 155 160  
 Val Val Ala Glu Glu Ile Ala Thr Phe Gly Leu Pro Ser Gly Ile Cys  
 165 170 175  
 Pro Ile Val Phe Val Phe Thr Gly Val Gly Asn Val Ser Gln Gly Ala  
 180 185 190  
 Gln Glu Ile Phe Lys Leu Leu Pro His Thr Phe Val Asp Ala Glu Lys  
 195 200 205  
 Leu Pro Glu Ile Phe Gln Ala Arg Asn Leu Ser Lys Gln Ser Gln Ser  
 210 215 220  
 Thr Lys Arg Val Phe Gln Leu Tyr Gly Cys Val Val Thr Ser Arg Asp  
 225 230 235 240  
 Ile Val Ser His Lys Asp Pro Thr Arg Gln Phe Asp Lys Gly Asp Tyr  
 245 250 255  
 Tyr Ala His Pro Glu His Tyr Thr Pro Val Phe His Glu Arg Ile Ala  
 260 265 270  
 Pro Tyr Ala Ser Val Ile Val Asn Cys Met Tyr Trp Glu Lys Arg Phe  
 275 280 285  
 Pro Pro Leu Leu Asn Met Asp Gln Leu Gln Gln Leu Met Glu Thr Gly  
 290 295 300  
 Cys Pro Leu Val Gly Val Cys Asp Ile Thr Cys Asp Ile Gly Gly Ser  
 305 310 315 320  
 Ile Glu Phe Ile Asn Lys Ser Thr Ser Ile Glu Arg Pro Phe Phe Arg  
 325 330 335  
 Tyr Asp Pro Ser Lys Asn Ser Tyr His Asp Asp Met Glu Gly Ala Gly  
 340 345 350  
 Val Val Cys Leu Ala Val Asp Ile Leu Pro Thr Glu Phe Ser Lys Glu  
 355 360 365  
 Ala Ser Gln His Phe Gly Asn Ile Leu Ser Arg Leu Val Ala Ser Leu  
 370 375 380

Ala Ser Val Lys Gln Pro Ala Glu Leu Pro Ser Tyr Leu Arg Arg Ala  
385 390 395 400

Cys Ile Ala His Ala Gly Arg Leu Thr Pro Leu Tyr Glu Tyr Ile Pro  
405 410 415

Arg Met Arg Asn Thr Met Ile Asp Leu Ala Pro Ala Lys Thr Asn Pro  
420 425 430

Leu Pro Asp Lys Lys Tyr Ser Thr Leu Val Ser Leu Ser Gly His Leu  
435 440 445

Phe Asp Lys Phe Leu Ile Asn Glu Ala Leu Asp Ile Ile Glu Thr Ala  
450 455 460

Gly Gly Ser Phe His Leu Val Arg Cys Glu Val Gly Gln Ser Thr Asp  
465 470 475 480

Asp Met Ser Tyr Ser Glu Leu Glu Val Gly Ala Asp Asp Thr Ala Thr  
485 490 495

Leu Asp Lys Ile Ile Asp Ser Leu Thr Ser Leu Ala Asn Glu His Gly  
500 505 510

Gly Asp His Asp Ala Gly Gln Glu Ile Glu Leu Ala Leu Lys Ile Gly  
515 520 525

Lys Val Asn Glu Tyr Glu Thr Asp Val Thr Ile Asp Lys Gly Gly Pro  
530 535 540

Lys Ile Leu Ile Leu Gly Ala Gly Arg Val Cys Arg Pro Ala Ala Glu  
545 550 555 560

Phe Leu Ala Ser Tyr Pro Asp Ile Cys Thr Tyr Gly Val Asp Asp His  
565 570 575

Asp Ala Asp Gln Ile His Val Ile Val Ala Ser Leu Tyr Gln Lys Asp  
580 585 590

Ala Glu Glu Thr Val Asp Gly Ile Glu Asn Thr Thr Ala Thr Gln Leu  
595 600 605

Asp Val Ala Asp Ile Gly Ser Leu Ser Asp Leu Val Ser Gln Val Glu  
610 615 620

Val Val Ile Ser Leu Leu Pro Ala Ser Phe His Ala Ala Ile Ala Gly  
625 630 635 640

(2) INFORMATION FOR SEQ ID NO:125:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 720 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Oryza sativa*

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..720
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 215
  - (D) OTHER INFORMATION: /label= unknown
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 678
  - (D) OTHER INFORMATION: /label= unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

```

GTTTAAACAT CTTTCCAATC TTGTTTCTCA GGTGGAAGTA GTAGTTAGCT TGCTGCCTGC      60
CAGTTTTTCAT GCTGCCATAG CAAGAGTATG CATAGAGATG AAGAAGCACT TGGTCACTGC      120
AAGCTATGTT GATGAGTCCA TGTCAAAGTT GGAACAATCT GCAGAAGGTG CTGGTGTAAC      180
TATTCTCTGT GAAATGGGCC TGGATCCTGG CATANATCAT ATGATGTCAA TGAAGATGAT      240
TGACGAAGCA CATTCACGGA AGGGGAAAAT AAAGTCATTT ACATCCTTTT GTGGAGGACT      300
TCCATCTCCA GCTTCTGCAA ACAATCCACT TGCTTATAAG TTCAGTTGGA GTCCAGCTGG      360
TGCCATCCGT GCAGGGAGAA ACCCTGCTGT CTACAAATTT CATGGAGAAA TCATCCATGT      420
AGATGGTGAT AAATTGTATG AATCCGCAAA GAGGCTCAGA TTACMAGAAC TTCCAGCTTT      480
TGCACTGGAA CACTTGCCAA ACCGGAATTC CTTGATGTAT GGAGACCTGT ATGGGATCTC      540
CAAAGAAGCA TCTACTGTGT ACAGGGCTAC TCTTCGTTAT GAAGGATTTA ATGAGATAAT      600
GGCAACCTTC GCGAAAATTG GGTTTTTTGA TGCTGCAAGT CATCCACTGT TGCAACAAAC      660
TACTCGCCCT ACATACANGG ATTTCTGTGT GAACCCTCAA TGCTTGTACA TCTCCAAAAC      720

```

(2) INFORMATION FOR SEQ ID NO:126:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 239 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein.
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Oryza sativa

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

Phe Lys His Leu Ser Asn Leu Val Ser Gln Val Glu Val Val Val Ser
1           5           10           15
Leu Leu Pro Ala Ser Phe His Ala Ala Ile Ala Arg Val Cys Ile Glu

```



CTGCTGTTGC TCCAGAACAA GATCCAAAAG AAAGGAGTGA TCAGGCCTCT GGAACCTGAA 60  
 ATTTACATTC CAGCGTTGGA GATCTTGGAG TCATCGGGTA TCAAGCTGGC GGAGAGAGTG 120  
 GAGACCTGAG AATCGGACCC AATATGTATA ATGTAGCATG GTGGTAGCTT CTCTATATAT 180  
 ATGCTTCAGT GAATAATTGA TTTGCCGTTG TGTGGTAATT AAGCAATGCC CGCTAATAAA 240  
 TTGTACCGTA GAAGTCCTTC TATGTACATC CGTATCAAAA AATAAAAAAA GCATCGATTA 300  
 GCTTGAAT 308

(2) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 

- (A) ORGANISM: *Oryza sativa*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Leu Leu Leu Leu Gln Asn Lys Ile Gln Lys Lys Gly Val Ile Arg Pro  
 1 5 10 15  
 Leu Glu Pro Glu Ile Tyr Ile Pro Ala Leu Glu Ile Leu Glu Ser Ser  
 20 25 30  
 Gly Ile Lys Leu Ala Glu Arg Val Glu Thr  
 35 40

(2) INFORMATION FOR SEQ ID NO:129:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 429 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 

- (A) ORGANISM: *Triticum aestivum*

(ix) FEATURE:
 

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..252

(ix) FEATURE:
 

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 172

(D) OTHER INFORMATION: /label= unknown

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 186  
(D) OTHER INFORMATION: /label= unknown

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 331  
(D) OTHER INFORMATION: /label= unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

```
TACCCCGACG GGGACCCAC CGAGAAGCAC CAAGCGACGC TGCTGGAGTT CGGAAAGACC 60
GAGAACGGCA GGCCACCAC CGCCATGGCC CTCACCGTTG GGTACCGGC AGCGATAGGA 120
GCCCTGCTCT TGCTCCAGAA CAAGGTCCAG AGGAAAGGGG TGATCCGGCC TNTGGAACCG 180
GAGATNTACA TCCCTGCGCT GGAGATCTTG GAAGCGTCGG GCATCAAGCT GATCGAGAGA 240
GTGGAGACCT GAGGATGTCA GGATGGGATG AGAATCTATC GAGTATATAT GCTGCAGCAA 300
CAGAGGCAGT GAGTAAATAA AATGATGATT NTCGCCGTTG TAAGTAAAAT GAGTGGACTG 360
TATGTATGTA TGTGACTATC TATTGTACTA CATATATACC AAATCTGTCT CCGGTTGATT 420
CTGTTGGTG 429
```

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 83 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Triticum aestivum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

```
Tyr Pro Asp Gly Asp Pro Thr Glu Lys His Gln Ala Thr Leu Leu Glu
1          5          10
Phe Gly Lys Thr Glu Asn Gly Arg Pro Thr Thr Ala Met Ala Leu Thr
20        25        30
Val Gly Val Pro Ala Ala Ile Gly Ala Leu Leu Leu Leu Gln Asn Lys
35        40        45
Val Gln Arg Lys Gly Val Ile Arg Pro Xaa Glu Pro Glu Xaa Tyr Ile
50        55        60
Pro Ala Leu Glu Ile Leu Glu Ala Ser Gly Ile Lys Leu Ile Glu Arg
65        70        75        80
```



Val Glu Thr

(2) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1449 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

```
ATGACGAAAA AATCAGGTGT TTTGATTCTT GGTGCTGGAC GTGTGTGTCTG CCCAGCTGCT    60
GATTTCTTAG CTTCAGTTAG AACCATTTTCG TCACAGCAAT GGTACAAAAC ATATTTTCGGA    120
GCAGACTCTG AAGAGAAAAC AGATGTTTCAT GTGATTGTCTG CGTCTCTGTA TCTTAAGGAT    180
GCCAAAGAGA CGGTTGAAGG TATTTTCAGAT GTAGAAGCAG TTCGGCTAGA TGTATCTGAT    240
AGTGAAAGTC TCCTTAAGTA TGTTTCTCAG GTTGATGTTG TCCTAAGTTT ATTACCTGCA    300
AGTTGTCATG CTGTTGTAGC AAAGACATGC ATTGAGCTGA AGAAGCATCT CGTCACTGCT    360
AGCTATGTTG ATGATGAAAC GTCCATGTTA CATGAGAAGG CTAAGAGTGC TGGGATAACG    420
ATTCTAGGCG AAATGGGACT GGACCCTGGA ATCGATCACA TGATGGCGAT GAAAATGATC    480
AACGATGCTC ATATCAAAAA AGGGAAAGTG AAGTCTTTTA CCTCTTATTG TGGAGGGCTT    540
CCCTCTCCTG CTGCAGCAAA TAATCCATTA GCATATAAAT TTAGCTGGAA CCCTGCTGGA    600
GCAATTTCGAG CTGGTCAAAA CCCCGCCAAA TACAAAAGCA ACGGCGACAT AATACATGTT    660
GATGGGAAGA ATCTCTATGA TTCCGCGGCA AGATTCCGAG TACCTAATCT TCCAGCTTTT    720
GCATTGGAGT GTTTTCCAAA TCGTGACTCC TTGGTTTACG GGGAACATTA TGGCATCGAG    780
AGCGAAGCAA CAACGATATT TCGTGGAACA CTCAGATATG AAGGGTTTAG TATGATAATG    840
GCAACACTTT CGAAACTTGG ATTCTTTGAC AGTGAAGCAA ATCAAGTACT CTCCACTGGA    900
AAGAGGATTA CGTTTGGTGC TCTTTTAAGT AACATTCTAA ATAAGGATGC AGACAATGAA    960
TCAGAGCCCC TAGCGGGAGA AGAAGAGATA AGCAAGAGAA TTATCAAGCT TGGACATTCC   1020
AAGGAGACTG CAGCCAAAGC TGCCAAAACA ATTGTATTCT TGGGGTTCAA CGAAGAGAGG   1080
GAGGTTCAT  CACTGTGTAA AAGCGTATTT GATGCAACTT GTTACCTAAT GGAAGAGAAA   1140
CTAGCTTATT CCGGAAATGA ACAGGACATG GTGCTTTTGC ATCACGAAGT AGAAGTGGA   1200
TTCCTTGAAA GCAAACGTAT AGAGAAGCAC ACTGCGACTC TTTTGGAATT CGGGGACATC   1260
AAGAATGGAC AAACAACAAC CGCTATGGCC AAGACTGTTG GGATCCCTGC AGCCATTGGA   1320
```

GCTCTGGTGT TAATTGAAGA CAAGATCAAG ACAAGAGGAG TCTTAAGGCC TCTCGAAGCA 1380  
 GAGGTGTATT TGCCAGCTTT GGATATATTG CAAGCATATG GTATAAAGCT GATGGAGAAG 1440  
 GCAGAATGA 1449

(2) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 482 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Met Thr Lys Lys Ser Gly Val Leu Ile Leu Gly Ala Gly Arg Val Cys  
 1 5 10 15  
 Arg Pro Ala Ala Asp Phe Leu Ala Ser Val Arg Thr Ile Ser Ser Gln  
 20 25 30  
 Gln Trp Tyr Lys Thr Tyr Phe Gly Ala Asp Ser Glu Glu Lys Thr Asp  
 35 40 45  
 Val His Val Ile Val Ala Ser Leu Tyr Leu Lys Asp Ala Lys Glu Thr  
 50 55 60  
 Val Glu Gly Ile Ser Asp Val Glu Ala Val Arg Leu Asp Val Ser Asp  
 65 70 75 80  
 Ser Glu Ser Leu Leu Lys Tyr Val Ser Gln Val Asp Val Val Leu Ser  
 85 90 95  
 Leu Leu Pro Ala Ser Cys His Ala Val Val Ala Lys Thr Cys Ile Glu  
 100 105 110  
 Leu Lys Lys His Leu Val Thr Ala Ser Tyr Val Asp Asp Glu Thr Ser  
 115 120 125  
 Met Leu His Glu Lys Ala Lys Ser Ala Gly Ile Thr Ile Leu Gly Glu  
 130 135 140  
 Met Gly Leu Asp Pro Gly Ile Asp His Met Met Ala Met Lys Met Ile  
 145 150 155 160  
 Asn Asp Ala His Ile Lys Lys Gly Lys Val Lys Ser Phe Thr Ser Tyr  
 165 170 175  
 Cys Gly Gly Leu Pro Ser Pro Ala Ala Ala Asn Asn Pro Leu Ala Tyr  
 180 185 190  
 Lys Phe Ser Trp Asn Pro Ala Gly Ala Ile Arg Ala Gly Gln Asn Pro  
 195 200 205

Ala Lys Tyr Lys Ser Asn Gly Asp Ile Ile His Val Asp Gly Lys Asn  
210 215 220

Leu Tyr Asp Ser Ala Ala Arg Phe Arg Val Pro Asn Leu Pro Ala Phe  
225 230 235 240

Ala Leu Glu Cys Phe Pro Asn Arg Asp Ser Leu Val Tyr Gly Glu His  
245 250 255

Tyr Gly Ile Glu Ser Glu Ala Thr Thr Ile Phe Arg Gly Thr Leu Arg  
260 265 270

Tyr Glu Gly Phe Ser Met Ile Met Ala Thr Leu Ser Lys Leu Gly Phe  
275 280 285

Phe Asp Ser Glu Ala Asn Gln Val Leu Ser Thr Gly Lys Arg Ile Thr  
290 295 300

Phe Gly Ala Leu Leu Ser Asn Ile Leu Asn Lys Asp Ala Asp Asn Glu  
305 310 315 320

Ser Glu Pro Leu Ala Gly Glu Glu Glu Ile Ser Lys Arg Ile Ile Lys  
325 330 335

Leu Gly His Ser Lys Glu Thr Ala Ala Lys Ala Ala Lys Thr Ile Val  
340 345 350

Phe Leu Gly Phe Asn Glu Glu Arg Glu Val Pro Ser Leu Cys Lys Ser  
355 360 365

Val Phe Asp Ala Thr Cys Tyr Leu Met Glu Glu Lys Leu Ala Tyr Ser  
370 375 380

Gly Asn Glu Gln Asp Met Val Leu Leu His His Glu Val Glu Val Glu  
385 390 395 400

Phe Leu Glu Ser Lys Arg Ile Glu Lys His Thr Ala Thr Leu Leu Glu  
405 410 415

Phe Gly Asp Ile Lys Asn Gly Gln Thr Thr Thr Ala Met Ala Lys Thr  
420 425 430

Val Gly Ile Pro Ala Ala Ile Gly Ala Leu Val Leu Ile Glu Asp Lys  
435 440 445

Ile Lys Thr Arg Gly Val Leu Arg Pro Leu Glu Ala Glu Val Tyr Leu  
450 455 460

Pro Ala Leu Asp Ile Leu Gln Ala Tyr Gly Ile Lys Leu Met Glu Lys  
465 470 475 480

Ala Glu